On the Optimality of Linear Precoding for Secrecy in the MIMO Broadcast Channel

S. Ali. A. Fakoorian and A. Lee Swindlehurst

Abstract

We study the optimality of linear precoding for the two-receiver multiple-input multiple-output (MIMO) Gaussian broadcast channel (BC) with confidential messages. Secret dirty-paper coding (S-DPC) is optimal under an input covariance constraint, but there is no computable secrecy capacity expression for the general MIMO case under an average power constraint. In principle, for this case, the secrecy capacity region could be found through an exhaustive search over the set of all possible matrix power constraints. Clearly, this search, coupled with the complexity of dirty-paper encoding and decoding, motivates the consideration of low complexity linear precoding as an alternative. We prove that for a two-user MIMO Gaussian BC under an input covariance constraint, linear precoding is optimal and achieves the same secrecy rate region as S-DPC if the input covariance constraint satisfies a specific condition, and we characterize the corresponding optimal linear precoders. We then use this result to derive a closed-form sub-optimal algorithm based on linear precoding for an average power constraint. Numerical results indicate that the secrecy rate region achieved by this algorithm is close to that obtained by the optimal S-DPC approach with a search over all suitable input covariance matrices.

I. INTRODUCTION

The work of Wyner [1] led to the development of the notion of secrecy capacity, which quantifies the maximum rate at which a transmitter can reliably send a secret message to a receiver, without an eavesdropper being able to decode it. More recently, researchers have considered secrecy for the two-user broadcast channel, where each receiver acts as an eavesdropper for the independent message transmitted to the other. This problem was addressed in [2], where inner and outer bounds for the secrecy capacity region were established. Further work in [3] studied the multiple-input single-output (MISO) Gaussian case, and [4] considered the general MIMO Gaussian case. It was shown in [4] that, under an input covariance constraint, both confidential

messages can be simultaneously communicated at their respective maximum secrecy rates, where the achievablity is obtained using secret dirty-paper coding (S-DPC). However, under an average power constraint, a computable secrecy capacity expression for the general MIMO case has not yet been derived. In principle, the secrecy capacity for this case could be found by an exhaustive search over the set of all input covariance matrices that satisfy the average power constraint [4]. Clearly, the complexity associated with such a search and the implementation of dirty-paper encoding and decoding make such an approach prohibitive except for very simple scenarios, and motivates the study of simpler techniques based on linear precoding.

While low-complexity linear transmission techniques have been extensively investigated for the broadcast channel (BC) without secrecy constraints, e.g., [5]-[7], there has been relatively little work on considering secrecy in the design of linear precoders for the BC case. In [8], we considered linear precoders for the MIMO Gaussian broadcast channel with confidential messages based on the generalized singular value decomposition (GSVD) [9], [10]. It was shown numerically in [8] that, with an optimal allocation of power for the GSVD-based precoder, the achievable secrecy rate is very close to the secrecy capacity region.

In this paper, we show that for a two-user MIMO Gaussian BC with arbitrary numbers of antennas at each node and under an input covariance constraint, linear precoding is optimal and achieves the same secrecy rate region as S-DPC for certain input covariance constraints, and we derive an expression for the optimal precoders in these scenarios. We then use this result to develop a sub-optimal closed-form algorithm for calculating linear precoders for the case of average power constraints. Our numerical results indicate that the secrecy rate region achieved by this algorithm is close to that obtained by the optimal S-DPC approach with a search over all suitable input covariance matrices.

In Section II, we describe the model for the MIMO Gaussian broadcast channel with confidential messages and the optimal S-DPC scheme, proposed in [4]. In Section III, we consider a general MIMO broadcast channel under a matrix covariance constraint, we derive the conditions under which linear precoding is optimal and achieves the same secrecy rate region as S-DPC, and we find the corresponding optimal precoders. We then present our sub-optimal algorithm for designing linear precoders for the case of an average power constraint in Section IV, followed

by numerical examples in Section V. Section VI concludes the paper.

Notation: Vector-valued random variables are written with non-boldface uppercase letters (*e.g.,* X), while the corresponding non-boldface lowercase letter ($\mathbf{x}$) denotes a specific realization of the random variable. Scalar variables are written with non-boldface (lowercase or uppercase) letters. The Hermitian (i.e., conjugate) transpose is denoted by $(.)^{H}$, the matrix trace by $\mathrm{Tr}(.)$, and $\mathbf{I}$ indicates an identity matrix. The inequality $\mathbf{A} \succ \mathbf{B}$ ($\mathbf{A} \succeq \mathbf{B}$) means that $\mathbf{A} - \mathbf{B}$ is Hermitian positive (semi-)definite. Mutual information between the random variables A and B is denoted by $I(A; B)$, $\mathbb{E}$ is the expectation operator, and $\mathcal{CN}(0, \sigma^2)$ represents the complex circularly symmetric Gaussian distribution with zero mean and variance σ^2.

II. BROADCAST CHANNEL AND S-DPC

We consider a two-receiver multiple-antenna Gaussian broadcast channel with confidential messages, where the transmitter, receiver 1 and receiver 2 possess n_t, m_1, and m_2 antennas, respectively. The transmitter has two independent confidential messages, W_1 and W_2, where W_1 is intended for receiver 1 but needs to be kept secret from receiver 2, and W_2 is intended for receiver 2 but needs to be kept secret from receiver 1 [4].

The signals at each receiver can be written as:

$$
\begin{aligned}
\mathbf{y}_1 &= \mathbf{H}\mathbf{x} + \mathbf{z}_1 \\
\mathbf{y}_2 &= \mathbf{G}\mathbf{x} + \mathbf{z}_2
\end{aligned}
\tag{1}
$$

where $\mathbf{x}$ is the $n_t \times 1$ transmitted signal, and $\mathbf{z}_i \in \mathbb{C}^{m_i \times 1}$ is white Gaussian noise at receiver i with independent and identically distributed entries drawn from $\mathcal{CN}(0, 1)$. The channel matrices $\mathbf{H} \in \mathbb{C}^{m_1 \times n_t}$ and $\mathbf{G} \in \mathbb{C}^{m_2 \times n_t}$ are assumed to be unrelated to each other, and known at all three nodes. The transmitted signal is subject to an average power constraint when

$$
\mathrm{Tr}(\mathbb{E}\{XX^H\}) = \mathrm{Tr}(\mathbf{Q}) \leq P_t
\tag{2}
$$

for some scalar P_t, or it is subject to a matrix power constraint when [4], [11]:

$$
\mathbb{E}\{XX^H\} = \mathbf{Q} \preceq \mathbf{S}
\tag{3}
$$

where $\mathbf{Q}$ is the transmit covariance matrix, and $\mathbf{S} \succeq 0$. Compared with the average power constraint, (3) is rather precise and inflexible, although for example it does allow for the incorporation of per-antenna power constraints as a special case.

It was shown in [2] that for any jointly distributed (V_1, V_2, X) such that $(V_1, V_2) \rightarrow X \rightarrow (Y_1, Y_2)$ forms a Markov chain and the power constraint over X is satisfied, the secrecy rate pair (R_1, R_2) given by

$$R_1 = I(V_1; Y_1) - I(V_1; V_2, Y_2)$$
$$R_2 = I(V_2; Y_2) - I(V_2; V_1, Y_1) \tag{4}$$

is achievable for the MIMO Gaussian broadcast channel given by (1), where the auxiliary variables V_1 and V_2 represent the precoding signals for the confidential messages W_1 and W_2, respectively [4]. In [2], the achievablity of the rate pair (4) was proved.

Liu *et al.* [4] analyzed the above secret communication problem under the matrix power-covariance constraint (3). They showed that the secrecy capacity region $\mathcal{C}_s(\mathbf{H}, \mathbf{G}, \mathbf{S})$ is rectangular. This interesting result implies that under the matrix power constraint, both confidential messages W_1 and W_2 can be *simultaneously* transmitted at their respective maximal secrecy rates, as if over two separate MIMO Gaussian wiretap channels. To prove this result, Liu *et al.* showed that the secrecy capacity of the MIMO Gaussian wiretap channel can also be achieved via a coding scheme that uses artificial noise and random binning [4, Theorem 2].

Under the matrix power constraint (3), the achievablity of the optimal corner point (R_1^*, R_2^*) given by [4, Theorem 1]

$$R_1^* = \max_{0 \preceq \mathbf{K}_t \preceq \mathbf{S}} \log \left| \mathbf{H}\mathbf{K}_t\mathbf{H}^H + \mathbf{I} \right| - \log \left| \mathbf{G}\mathbf{K}_t\mathbf{G}^H + \mathbf{I} \right|$$
$$R_2^* = \log \left| \frac{\mathbf{G}\mathbf{S}\mathbf{G}^H + \mathbf{I}}{\mathbf{H}\mathbf{S}\mathbf{H}^H + \mathbf{I}} \right| + R_1^* \tag{5}$$

is obtained using dirty-paper coding based on double binning, or as referred to in [4], secret dirty paper coding (S-DPC). More precisely, let $\mathbf{K}_t^* \succeq \underline{0}$ maximize (5), and let

$$V_1 = U_1 + \mathbf{F}U_2 \qquad V_2 = U_2 \qquad X = U_1 + U_2 \ , \tag{6}$$

where U_1 and U_2 are two independent Gaussian vectors with zero means and covariance matrices $\mathbf{K}_t^*$ and $\mathbf{S} - \mathbf{K}_t^*$, respectively, and the precoding matrix $\mathbf{F}$ is defined as $\mathbf{F} = \mathbf{K}_t^* \mathbf{H}^H \left(\mathbf{H} \mathbf{K}_t^* \mathbf{H}^H + \mathbf{I} \right)^{-1} \mathbf{H}$. One can easily confirm the achievablity of the corner point (R_1^*, R_2^*) by evaluating (4) for the above random variables and noting that in (1), $X = U_1 + U_2$. Note that under the matrix power constraint $\mathbf{S}$, the input covariance matrix that achieves the corner point in the secrecy capacity region satisfies $\mathbf{Q} = \mathbf{S}$ [4].

The matrix $\mathbf{K}_t$ that maximizes (5) is given by [4], [11]

$$\mathbf{K}_t^* = \mathbf{S}^{\frac{1}{2}} \mathbf{C} \begin{bmatrix} (\mathbf{C}_1^H \mathbf{C}_1)^{-1} & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \mathbf{C}^H \mathbf{S}^{\frac{1}{2}} \tag{7}$$

where $\mathbf{C} = [\mathbf{C}_1 \ \mathbf{C}_2]$ is an invertible[1] generalized eigenvector matrix of the pencil

$$\left(\mathbf{S}^{\frac{1}{2}} \mathbf{H}^H \mathbf{H} \mathbf{S}^{\frac{1}{2}} + \mathbf{I} \,, \ \mathbf{S}^{\frac{1}{2}} \mathbf{G}^H \mathbf{G} \mathbf{S}^{\frac{1}{2}} + \mathbf{I} \right) \tag{8}$$

satisfying [12]

$$\begin{aligned} \mathbf{C}^H \left[\mathbf{S}^{\frac{1}{2}} \mathbf{H}^H \mathbf{H} \mathbf{S}^{\frac{1}{2}} + \mathbf{I} \right] \mathbf{C} &= \mathbf{\Lambda} \\ \mathbf{C}^H \left[\mathbf{S}^{\frac{1}{2}} \mathbf{G}^H \mathbf{G} \mathbf{S}^{\frac{1}{2}} + \mathbf{I} \right] \mathbf{C} &= \mathbf{I} \,, \end{aligned} \tag{9}$$

where $\mathbf{\Lambda} = \mathrm{diag}\{\lambda_1, ..., \lambda_{n_t}\} \succ \underline{0}$ contains the generalized eigenvalues sorted without loss of generality such that

$$\lambda_1 \geq ... \geq \lambda_b > 1 \geq \lambda_{b+1} \geq ... \geq \lambda_{n_t} > 0 \,.$$

The quantity b denotes the number of generalized eigenvalues greater than one ($0 \leq b \leq n_t$), and defines the following matrix partitions:

$$\mathbf{\Lambda} = \begin{bmatrix} \mathbf{\Lambda}_1 & 0 \\ 0 & \mathbf{\Lambda}_2 \end{bmatrix} \qquad \mathbf{C} = [\mathbf{C}_1 \ \mathbf{C}_2] \,, \tag{10}$$

where $\mathbf{\Lambda}_1 = \mathrm{diag}\{\lambda_1, ..., \lambda_b\}$, $\mathbf{\Lambda}_2 = \mathrm{diag}\{\lambda_{b+1}, ..., \lambda_{n_t}\}$, $\mathbf{C}_1$ contains the b generalized eigenvectors corresponding to $\mathbf{\Lambda}_1$ and $\mathbf{C}_2$ the $(n_t - b)$ generalized eigenvectors corresponding to $\mathbf{\Lambda}_2$.

---

[1]Note that $\mathbf{C}$ is invertible since both components of the pencil (8) are positive definite.

Now, by applying (7) in (5), the corner rate pair (R_1^*, R_2^*) can be calculated as ([4, Theorem 3])

$$R_1^* = \log |\mathbf{\Lambda}_1|$$
$$R_2^* = -\log |\mathbf{\Lambda}_2| \ . \tag{11}$$

For the average power constraint in (2), there is no computable secrecy capacity expression for the general MIMO case. In principle the secrecy capacity region for the average power constraint, $\mathcal{C}_s(\mathbf{H}, \mathbf{G}, P_t)$, could be found through an exhaustive search over all suitable matrix power constraints [4],[13, Lemma 1]:

$$\mathcal{C}_s(\mathbf{H}, \mathbf{G}, P_t) = \bigcup_{\mathbf{S} \succeq 0, \mathrm{Tr}(\mathbf{S}) \leq P_t} \mathcal{C}_s(\mathbf{H}, \mathbf{G}, \mathbf{S}) \ . \tag{12}$$

For any given semidefinite $\mathbf{S}$, $\mathcal{C}_s(\mathbf{H}, \mathbf{G}, \mathbf{S})$ can be computed as given by (11). Then, the secrecy capacity region $\mathcal{C}_s(\mathbf{H}, \mathbf{G}, P_t)$ is the convex hull of all of the obtained corner points using (11).

The complexity associated with such a search, as well as that required to implement dirty-paper encoding and decoding, are the main drawbacks of using S-DPC to find the secrecy capacity region $\mathcal{C}_s(\mathbf{H}, \mathbf{G}, P_t)$ for the average power constraint. This makes linear precoding (beamforming) techniques an attractive alternative because of their simplicity. To address the performance achievable with linear precoding, we first describe the conditions under which linear precoding is optimal in attaining the same secrecy rate region that is achievable via S-DPC, when the broadcast channel is under an input covariance constraint. In particular, in the next section we show that this equivalence holds for matrix power constraints that satisfy a certain property, and we derive the linear precoders that achieve optimal performance. Section IV then uses these results to derive a sub-optimal algorithm for the case of the average power constraint.

III. OPTIMALITY OF LINEAR PRECODING FOR BC SECRECY

In this section we answer the following questions:

(a) For a given general MIMO Gaussian BC described by (1), where each node has an arbitrary number of antennas and the channel input is under the covariance constraint (3), is there any $\mathbf{S} \succeq \mathbf{0}$ for which linear precoding can attain the secrecy capacity region?

(b) If yes, how can such $\mathbf{S}$ be described?

(c) For such $\mathbf{S}$, what is the optimal linear precoder that allows the rectangular S-DPC capacity region given by (11) to be achieved?

(d) If $\mathbf{S}$ does not satisfy the condition for optimal linear precoding in (a), what is the worst-case loss in secrecy capacity incurred by using the linear precoding approach described in (b) anyway?

To begin, we give the following theorem as an answer to questions (a) and (b) above.

Theorem 1. Suppose the matrix power constraint $\mathbf{S} \succeq \underline{0}$ on the input covariance $\mathbf{Q}$ in (3) leads to generalized eigenvectors in (9) that satisfy $\mathrm{span}\{\mathbf{C}_1\} \perp \mathrm{span}\{\mathbf{C}_2\}$, i.e. $\mathbf{C}_1^H \mathbf{C}_2 = \underline{0}$. Then the secrecy capacity region $\mathcal{C}_s(\mathbf{H}, \mathbf{G}, \mathbf{S})$ can be achieved with $X = V_1 + V_2$, where V_1 and V_2 are *independent* Gaussian precoders respectively corresponding to W_1 and W_2, with zero means and covariance matrices $\mathbf{K}_t^*$ and $\mathbf{S} - \mathbf{K}_t^*$, with $\mathbf{K}_t^*$ defined in (7).

Proof: Recall that for any $\mathbf{S} \succeq \underline{0}$, the secrecy capacity region $\mathcal{C}_s(\mathbf{H}, \mathbf{G}, \mathbf{S})$ is rectangular, so we only need to show that when $\mathbf{C}_2^H \mathbf{C}_1 = \underline{0}$, the linear precoders V_1 and V_2 characterized in this theorem are capable of achieving the corner point (R_1^*, R_2^*) given by (11). From (4), the achievable secrecy rate R_1 is given by

$$R_1 = I(V_1; Y_1) - I(V_1; V_2) - I(V_1; Y_2|V_2) = I(V_1; Y_1) - I(V_1; Y_2|V_2) \tag{13}$$

$$= I(V_1; \mathbf{H}(V_1 + V_2) + Z_1) - I(V_1; \mathbf{G}(V_1 + V_2) + Z_2|V_2)$$

$$= I(V_1; \mathbf{H}(V_1 + V_2) + Z_1) - I(V_1; \mathbf{G}V_1 + Z_2) \tag{14}$$

$$= \log|\mathbf{\Lambda}_1| = R_1^*, \tag{15}$$

where (13) and the second part of (14) come from the fact that V_1 and V_2 are independent. Equation (15) is proved in Appendix A. One can similarly show that $R_2 = R_2^* = -\log|\mathbf{\Lambda}_2|$ is achievable to complete the proof. ∎

Theorem 1 shows that the secrecy capacity region corresponding to any $\mathbf{S}$ with orthogonal $\mathbf{C}_1$ and $\mathbf{C}_2$ can be achieved using either linear independent precoders V_1 and V_2, as defined in Theorem 1, or using the S-DPC approach, as given by (6). The next theorem expands on the answer to question (b) above, and also addresses (c). First however we present the following

lemma which holds for any $\mathbf{S} \succeq \underline{0}$.

Lemma 1. For a given BC under the matrix power constraint (3), for any $\mathbf{S} \succeq \underline{0}$ we have $\mathrm{rank}(\mathbf{C}_1) \leq m$, where m is the number of positive eigenvalues of the matrix $\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G}$.

Proof: Please see Appendix B. ∎

The following theorem presents a more specific condition on $\mathbf{S}$ that results in generalized eigenvectors that satisfy $\mathrm{span}\{\mathbf{C}_1\} \perp \mathrm{span}\{\mathbf{C}_2\}$.

Theorem 2. For any $\mathbf{S} \succeq \underline{0}$, the generalized eigenvectors $\mathbf{C}_1$ and $\mathbf{C}_2$ in (9) are orthogonal *iff* there exists a matrix $\mathbf{T} \in \mathbb{C}^{n_t \times n_t}$ such that $\mathbf{S} = \mathbf{T}\mathbf{T}^H$ and $\mathbf{T}$ simultaneously block diagonalizes $\mathbf{H}^H\mathbf{H}$ and $\mathbf{G}^H\mathbf{G}$:

$$\mathbf{T}^H\mathbf{H}^H\mathbf{H}\mathbf{T} = \begin{bmatrix} \mathbf{K_{H1}} & \underline{0} \\ \underline{0} & \mathbf{K_{H2}} \end{bmatrix} \qquad \mathbf{T}^H\mathbf{G}^H\mathbf{G}\mathbf{T} = \begin{bmatrix} \mathbf{K_{G1}} & \underline{0} \\ \underline{0} & \mathbf{K_{G2}} \end{bmatrix}, \qquad (16)$$

where the $m \times m$ matrices $\mathbf{K_{H1}} \succeq \underline{0}$ and $\mathbf{K_{G1}} \succeq \underline{0}$ satisfy $\mathbf{K_{H1}} \succeq \mathbf{K_{G1}}$ and $\mathbf{K_{H2}} \preceq \mathbf{K_{G2}}$.

Proof: The proof begins by noting that if $\mathbf{S} = \mathbf{T}\mathbf{T}^H$, then the pencil in (8) and

$$\left(\mathbf{T}^H\mathbf{H}^H\mathbf{H}\mathbf{T} + \mathbf{I} \,,\, \mathbf{T}^H\mathbf{G}^H\mathbf{G}\mathbf{T} + \mathbf{I}\right)$$

have exactly the same generalized eigenvalue matrix $\mathbf{\Lambda}$, and thus the same secrecy capacity regions. The remainder of the proof can be found in Appendix C. ∎

While algorithms exist to find $\mathbf{T}$ that jointly block diagonalizes $\mathbf{H}^H\mathbf{H}$ and $\mathbf{G}^H\mathbf{G}$ (see for example [14] and references therein), as mentioned in Appendix C only those $\mathbf{T}$ that lead to $\mathbf{K_{H1}} \succeq \mathbf{K_{G1}}$ and $\mathbf{K_{H2}} \preceq \mathbf{K_{G2}}$ are acceptable. Later, we will demonstrate that for any BC there are an infinite number of matrix constraints $\mathbf{S}$ that can achieve such a block diagonalization and hence allow for an optimal linear precoding solution.

To conclude this section, we now answer question (d) posed above. Define the projection matrices $\mathbf{P_{C_i}} = \mathbf{C}_i(\mathbf{C}_i^H\mathbf{C}_i)^{-1}\mathbf{C}_i^H$ and $\mathbf{P}_{\mathbf{C}_i}^\perp = \mathbf{I} - \mathbf{P_{C_i}}$, and note that in general, equation (7) is equivalent to $\mathbf{K}_t^* = \mathbf{S}^{\frac{1}{2}}\mathbf{P_{C_1}}\mathbf{S}^{\frac{1}{2}}$ and $\mathbf{S} - \mathbf{K}_t^* = \mathbf{S}^{\frac{1}{2}}\mathbf{P}_{\mathbf{C}_1}^\perp\mathbf{S}^{\frac{1}{2}}$. When $\mathbf{C}_1^H\mathbf{C}_2 = \underline{0}$, the optimal

covariance matrices for V_1 and V_2 also satisfy

$$\mathbf{K}_t^* = \mathbf{S}^{\frac{1}{2}} \mathbf{P}_{\mathbf{C}_2}^{\perp} \mathbf{S}^{\frac{1}{2}} \tag{17}$$

$$\mathbf{S} - \mathbf{K}_t^* = \mathbf{S}^{\frac{1}{2}} \mathbf{P}_{\mathbf{C}_2} \mathbf{S}^{\frac{1}{2}} . \tag{18}$$

The following theorem explains the loss in secrecy that results when linear precoding with these covariances is used for a matrix constraint $\mathbf{S}$ that does *not* satisfy $\mathbf{C}_1^H \mathbf{C}_2 = \underline{0}$.

Theorem 3. Assume a linear precoding scheme $X = V_1 + V_2$ for independent Gaussian precoders V_1 and V_2 with zero means and covariance matrices $\mathbf{K}_t = \mathbf{S}^{\frac{1}{2}} \mathbf{P}_{\mathbf{C}_2}^{\perp} \mathbf{S}^{\frac{1}{2}}$ and $\mathbf{S} - \mathbf{K}_t = \mathbf{S}^{\frac{1}{2}} \mathbf{P}_{\mathbf{C}_2} \mathbf{S}^{\frac{1}{2}}$, respectively. Also define $\mathbf{N} = \left(\mathbf{C}_2^H \mathbf{P}_{\mathbf{C}_1}^{\perp} \mathbf{C}_2 \right)^{-1} \mathbf{C}_2^H \mathbf{P}_{\mathbf{C}_1}^{\perp} \mathbf{P}_{\mathbf{C}_2}^{\perp} \mathbf{C}_1$. The loss in secrecy capacity that results from using this approach in the two-user BC is *at most* $\log \left| \mathbf{I} + \mathbf{N}^H \mathbf{N} \right|$ for each user. In particular, the following secrecy rate pair is achievable:

$$
\begin{aligned}
R_1 &= \max(0, \, R_1^* - \log \left| \mathbf{I} + \mathbf{N}^H \mathbf{N} \right|) = \max(0, \, \log |\mathbf{\Lambda}_1| - \log \left| \mathbf{I} + \mathbf{N}^H \mathbf{N} \right|) \\
R_2 &= \max(0, \, R_2^* - \log \left| \mathbf{I} + \mathbf{N}^H \mathbf{N} \right|) = \max(0, \, -\log |\mathbf{\Lambda}_2| - \log \left| \mathbf{I} + \mathbf{N}^H \mathbf{N} \right|) .
\end{aligned}
\tag{19}
$$

Proof: See Appendix D. ■

Remark 1. Note that if $\mathbf{C}_1$ and $\mathbf{C}_2$ are orthogonal, then $\mathbf{N} = \underline{0}$ and $R_i = R_i^*$, $i = 1, 2$, is achievable, as discussed in Theorem 1.

IV. SUB-OPTIMAL SOLUTIONS UNDER AN AVERAGE POWER CONSTRAINT

So far we have shown that if the broadcast channel (1) is under the matrix power constraint $\mathbf{S}$ (3), then linear precoding as defined by Theorem 1 is an optimal solution when $\mathbf{S}$ satisfies the condition described in Theorem 2. In the following we propose a suboptimal closed-form linear precoding scheme for the general MIMO Gaussian BC under the *average* power constraint (2), where as mentioned earlier there exists no optimal closed-form solution that characterizes the secrecy capacity region. We begin with some preliminary results, then we develop the algorithm for the general MIMO case, and finally we present an alternative algorithm specifically for the MISO case since it offers additional insight.

A. Preliminary Results

Remark 2. Suppose that the input covariance matrix $\mathbf{Q}$ leads to a point on the Pareto boundary of the secrecy capacity region given by (12) under the average power constraint (2). Then $\mathrm{Tr}(\mathbf{Q}) = P_t$ and $\mathbf{Q}$ cannot have any component in the nullspace of $\mathbf{H}^H\mathbf{H} + \mathbf{G}^H\mathbf{G}$, and thus $\mathcal{C}_s(\mathbf{H}, \mathbf{G}, P_t) = \mathcal{C}_s(\mathbf{H}_{eq}, \mathbf{G}_{eq}, P_t)$, where $\mathbf{H}_{eq} = \mathbf{H}\mathbf{U}_p$, $\mathbf{G}_{eq} = \mathbf{G}\mathbf{U}_p$ and $\mathbf{U}_p$ contains the singular vectors corresponding to the non-zero singular values of $\mathbf{H}^H\mathbf{H} + \mathbf{G}^H\mathbf{G}$.

According to Remark 2, we can assume without loss of generality that $\mathbf{H}^H\mathbf{H} + \mathbf{G}^H\mathbf{G}$ is full-rank; otherwise, we could replace $\mathbf{H}, \mathbf{G}$ with $\mathbf{H}_{eq}, \mathbf{G}_{eq}$ and have an equivalent problem where $\mathbf{H}_{eq}^H\mathbf{H}_{eq} + \mathbf{G}_{eq}^H\mathbf{G}_{eq}$ is full-rank and the secrecy capacity region is the same (in such a case, n_t would then represent the number of transmitted data streams rather than the number of antennas). With this result, we have the following lemma.

Lemma 2. Define

$$\mathbf{W} = (\mathbf{H}^H\mathbf{H} + \mathbf{G}^H\mathbf{G})^{-\frac{1}{2}} . \tag{20}$$

Then $\mathbf{W}\mathbf{H}^H\mathbf{H}\mathbf{W}$ and $\mathbf{W}\mathbf{G}^H\mathbf{G}\mathbf{W}$ commute and hence share the same set of eigenvectors:

$$\begin{aligned}
\mathbf{W}\mathbf{H}^H\mathbf{H}\mathbf{W} &= \boldsymbol{\Phi}_\mathbf{w}\boldsymbol{\Sigma}_1\boldsymbol{\Phi}_\mathbf{w}^H \\
\mathbf{W}\mathbf{G}^H\mathbf{G}\mathbf{W} &= \boldsymbol{\Phi}_\mathbf{w}\boldsymbol{\Sigma}_2\boldsymbol{\Phi}_\mathbf{w}^H ,
\end{aligned} \tag{21}$$

where $\boldsymbol{\Phi}_\mathbf{w}$ is the (unitary) matrix of eigenvectors and $\boldsymbol{\Sigma}_1 \succeq \underline{0}, \boldsymbol{\Sigma}_2 \succeq \underline{0}$ the corresponding eigenvalues.

Proof: See Appendix E. ∎

Without loss of generality, we assume that the columns of $\boldsymbol{\Phi}_\mathbf{w}$ are sorted such that the first ρ diagonal elements of $\boldsymbol{\Sigma}_1$ are greater than the first ρ diagonal elements of $\boldsymbol{\Sigma}_2$, and the last $n_t - \rho$ diagonal elements of $\boldsymbol{\Sigma}_1$ are less than or equal to those of $\boldsymbol{\Sigma}_2$. Recall from Lemma 1 that $0 \le \rho \le m$, where m is the number of positive eigenvalues of $\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G}$. Thus,

$$\boldsymbol{\Sigma}_1 = \begin{bmatrix} \boldsymbol{\Sigma}_{1\rho} & \underline{0} \\ \underline{0} & \boldsymbol{\Sigma}_{1\bar{\rho}} \end{bmatrix} \qquad \boldsymbol{\Sigma}_2 = \begin{bmatrix} \boldsymbol{\Sigma}_{2\rho} & \underline{0} \\ \underline{0} & \boldsymbol{\Sigma}_{2\bar{\rho}} \end{bmatrix} \tag{22}$$

where $\mathbf{\Sigma}_{i\rho}$ is $\rho \times \rho$, $\mathbf{\Sigma}_{i\bar{\rho}}$ is $(n_t - \rho) \times (n_t - \rho)$, $\mathbf{\Sigma}_{1\rho} \succ \mathbf{\Sigma}_{2\rho}$ and $\mathbf{\Sigma}_{1\bar{\rho}} \preceq \mathbf{\Sigma}_{2\bar{\rho}}$.

Now define

$$\mathbf{S_w} = \mathbf{T_w}\mathbf{T_w}^H = \mathbf{W}\mathbf{\Phi_w}\ \mathbf{P}\ \mathbf{\Phi_w}^H \mathbf{W} \tag{23}$$

$$\mathbf{T_w} = \mathbf{W}\mathbf{\Phi_w}\mathbf{P}^{\frac{1}{2}}\ , \tag{24}$$

where $\mathbf{W}$ and $\mathbf{\Phi_w}$ are given in Lemma 2 and $\mathbf{P} \succeq \underline{0}$ is any block-diagonal matrix partitioned in the same way as $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$. With these definitions, we see from (21) that $\mathbf{T_w}^H\mathbf{H}^H\mathbf{H}\mathbf{T_w}$ and $\mathbf{T_w}^H\mathbf{G}^H\mathbf{G}\mathbf{T_w}$ are block diagonal. Thus, from Theorem 2, a BC with the matrix power constraint $\mathbf{S_w} = \mathbf{T_w}^H\mathbf{T_w}$ leads to a matrix pencil $\left(\mathbf{S_w}^{\frac{1}{2}}\mathbf{H}^H\mathbf{H}\mathbf{S_w}^{\frac{1}{2}} + \mathbf{I}\ ,\ \mathbf{S_w}^{\frac{1}{2}}\mathbf{G}^H\mathbf{G}\mathbf{S_w}^{\frac{1}{2}} + \mathbf{I}\right)$ with generalized eigenvectors $\mathbf{C_w} = [\mathbf{C_{1w}}\ \ \mathbf{C_{2w}}]$ that satisfy $\mathbf{C_{1w}}^H\mathbf{C_{2w}} = \underline{0}$, where $\mathbf{C_{1w}}, \mathbf{C_{2w}}$ correspond to generalized eigenvalues that are larger or less-than-or-equal-to one, respectively.

Remark 3. Since the above result holds for any block-diagonal $\mathbf{P} \succeq 0$ with appropriate dimensions, then for every BC there are an infinite number of matrix power constraints $\mathbf{S_w}$ that achieve a block diagonalization and hence allow for an *optimal* linear precoding solution.

In the following, we restrict our attention to diagonal rather than block-diagonal matrices $\mathbf{P}$, for which a closed form solution can be derived. From Theorem 1, we have the following result.

Lemma 3. For any diagonal $\mathbf{P} \succeq 0$, the secrecy capacity of the broadcast channel in (1) under the matrix power constraint $\mathbf{S_w} = \mathbf{T_w}\mathbf{T_w}^H$ defined in (20)-(24) can be obtained by linear precoding. In particular,

$$X = \mathbf{W}\mathbf{\Phi_w}\begin{bmatrix} V_1' \\ V_2' \end{bmatrix} = V_1 + V_2 \tag{25}$$

where $V_1' \in \mathbb{C}^\rho$ and $V_2' \in \mathbb{C}^{n_t - \rho}$ are independent Gaussian random vectors with zero means and covariance matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ such that

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 & 0 \\ 0 & \mathbf{P}_2 \end{bmatrix}\ , \tag{26}$$

and as before V_1, V_2 represent independently encoded Gaussian codebook symbols corresponding

to the confidential messages W_1 and W_2, with zero means and covariances $\mathbf{K}_{t\mathbf{w}}^*$ and $\mathbf{S_w} - \mathbf{K}_{t\mathbf{w}}^*$ respectively given by

$$\mathbf{K}_{t\mathbf{w}}^* = \mathbf{W}\mathbf{\Phi_w} \begin{bmatrix} \mathbf{P}_1 & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \mathbf{\Phi_w^H W} \tag{27}$$

$$\mathbf{S_w} - \mathbf{K}_{t\mathbf{w}}^* = \mathbf{W}\mathbf{\Phi_w} \begin{bmatrix} \underline{0} & \underline{0} \\ \underline{0} & \mathbf{P}_2 \end{bmatrix} \mathbf{\Phi_w^H W} . \tag{28}$$

Proof: The matrix $\mathbf{T_w}$ simultaneously block diagonalizes $\mathbf{H}^H\mathbf{H}$ and $\mathbf{G}^H\mathbf{G}$, so by Theorems 1 and 2 we know that linear precoding can achieve the secrecy capacity region. The proof is completed in Appendix F by showing the equality in (25), and showing that (27) corresponds to the optimal covariance in (7). ∎

From the proof in Appendix F and (21)-(24), we see that under the matrix power constraint $\mathbf{S_w}$ given by (23) with diagonal $\mathbf{P}$, the general BC is transformed to an equivalent BC with a set of parallel independent subchannels between the transmitter and the receivers, and it suffices for the transmitter to use independent Gaussian codebooks across these subchannels. In particular, the diagonal entries of $\mathbf{P}_1$ and $\mathbf{P}_2$ represent the power assigned to these independent subchannels prior to application of the precoder $\mathbf{W}\mathbf{\Phi_w}$ in (24)[2]. From (25), the signals at the two receivers are given by

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{HW}\mathbf{\Phi_w} \begin{bmatrix} \mathbf{v}_1' \\ \mathbf{v}_2' \end{bmatrix} + \mathbf{z}_1 \\ &= \mathbf{\Gamma}_1\mathbf{\Sigma}_1 \begin{bmatrix} \mathbf{v}_1' \\ \mathbf{v}_2' \end{bmatrix} + \mathbf{z}_1 \\ &= \mathbf{\Gamma}_1 \begin{bmatrix} \mathbf{\Sigma}_{1\rho}\mathbf{v}_1' \\ \mathbf{\Sigma}_{1\bar{\rho}}\mathbf{v}_2' \end{bmatrix} + \mathbf{z}_1 \\ \mathbf{y}_2 &= \mathbf{\Gamma}_2 \begin{bmatrix} \mathbf{\Sigma}_{2\rho}\mathbf{v}_1' \\ \mathbf{\Sigma}_{2\bar{\rho}}\mathbf{v}_2' \end{bmatrix} + \mathbf{z}_2 , \end{aligned}$$

[2]Note that the matrices $\mathbf{P}_1, \mathbf{P}_2$ do not represent the actual transmitted power, since the columns of $\mathbf{W}\mathbf{\Phi_w}$ are not unit-norm.

where $\mathbf{\Gamma}_1, \mathbf{\Gamma}_2$ are unitary. The confidential message for receiver 1 is thus transmitted with power loading $\mathbf{P}_1$ over those subchannels which are degraded for receiver 2 ($\mathbf{\Sigma}_{1\rho} \succ \mathbf{\Sigma}_{2\rho}$), while receiver 2's confidential message has power loading $\mathbf{P}_2$ over subchannels which are degraded for receiver 1 ($\mathbf{\Sigma}_{2\bar{\rho}} \succ \mathbf{\Sigma}_{1\bar{\rho}}$). Any subchannels for which the diagonal elements of $\mathbf{\Sigma}_{2\bar{\rho}}$ are equal to those of $\mathbf{\Sigma}_{1\bar{\rho}}$ are useless from the viewpoint of secret communication, but could be used to send common non-confidential messages.

From Theorem 1, the rectangular secrecy capacity region of the MIMO Gaussian BC (1) under the matrix power constraint $\mathbf{S_w}$ (23) is defined by the corner points

$$
\begin{aligned}
R_1^*(\mathbf{P}_1) &= \log|\mathbf{\Lambda}_{1\mathbf{w}}| = \log|\mathbf{I} + \mathbf{\Sigma}_{1\rho}\mathbf{P}_1| - \log|\mathbf{I} + \mathbf{\Sigma}_{2\rho}\mathbf{P}_1| \\
R_2^*(\mathbf{P}_2) &= -\log|\mathbf{\Lambda}_{2\mathbf{w}}| = \log|\mathbf{I} + \mathbf{\Sigma}_{2\bar{\rho}}\mathbf{P}_2| - \log|\mathbf{I} + \mathbf{\Sigma}_{1\bar{\rho}}\mathbf{P}_2| \ ,
\end{aligned}
\tag{29}
$$

where $\mathbf{\Lambda}_{i\mathbf{w}}$ is given by (81) in Appendix F. Note that we have explicitly written R_1^* as a function of the diagonal matrix $\mathbf{P}_1 \succeq \underline{0}$ to emphasize that $\mathbf{P}_1$ contains the only parameters that can be optimized for R_1^*. More precisely, since for a given matrix power constraint $\mathbf{S_w}$, $\mathbf{\Sigma}_{1\rho}$ and $\mathbf{\Sigma}_{2\rho}$ are channel dependent and thus fixed, as shown in (21)-(22). A similar description is also true for R_2^*.

B. Algorithm for the MIMO Case Under the Average Power Constraint

Here we propose our sub-optimal closed form solution based on linear precoding for the broadcast channel under the *average* power constraint (2). The goal is to find the diagonal matrix $\mathbf{P}$ in (23) that maximizes R_i^* in (29) for a given allocation of the transmit power to message W_i, and that satisfies the average power constraint[3]

$$
\begin{aligned}
\mathrm{Tr}(E\{XX^H\}) = \mathrm{Tr}(\mathbf{S_w}) &= \mathrm{Tr}(\mathbf{W}\mathbf{\Phi_w}\mathbf{P}\mathbf{\Phi_w}^H\mathbf{W}) \\
&= \mathrm{Tr}(\mathbf{\Phi_w}^H\mathbf{W}^2\mathbf{\Phi_w}\mathbf{P}) = \mathrm{Tr}\left(\mathbf{\Phi_w}^H(\mathbf{H}^H\mathbf{H} + \mathbf{G}^H\mathbf{G})^{-1}\mathbf{\Phi_w}\,\mathbf{P}\right) = P_t \ .
\end{aligned}
\tag{30}
$$

---

[3]Note that since we want to characterize the achievable secrecy rate points on the Pareto boundary, we use an equality constraint on the total power P_t in (30).

Noting that $\boldsymbol{\Phi_w}$ can be written as $\boldsymbol{\Phi_w} = [\boldsymbol{\Phi_{1w}} \ \boldsymbol{\Phi_{2w}}]$, where $\boldsymbol{\Phi_{1w}}$ is a $n_t \times \rho$ submatrix corresponding to the eigenvalues in $\boldsymbol{\Sigma_{1\rho}}$, (30) can be rewritten as

$$
\begin{aligned}
\mathrm{Tr}(E\{XX^H\}) &= \mathrm{Tr}\left(\boldsymbol{\Phi_w^H}(\mathbf{H}^H\mathbf{H} + \mathbf{G}^H\mathbf{G})^{-1}\boldsymbol{\Phi_w}\ \mathbf{P}\right) \\
&= \mathrm{Tr}\left(\boldsymbol{\Phi_{1w}^H}(\mathbf{H}^H\mathbf{H} + \mathbf{G}^H\mathbf{G})^{-1}\boldsymbol{\Phi_{1w}}\ \mathbf{P}_1\right) + \mathrm{Tr}\left(\boldsymbol{\Phi_{2w}^H}(\mathbf{H}^H\mathbf{H} + \mathbf{G}^H\mathbf{G})^{-1}\boldsymbol{\Phi_{2w}}\ \mathbf{P}_2\right) \\
&= \mathrm{Tr}\left(\mathbf{A}_1\mathbf{P}_1\right) + \mathrm{Tr}\left(\mathbf{A}_2\mathbf{P}_2\right) = P_t
\end{aligned}
\tag{31}
$$

where we defined positive definite matrices $\mathbf{A}_i = \boldsymbol{\Phi_{iw}^H}(\mathbf{H}^H\mathbf{H} + \mathbf{G}^H\mathbf{G})^{-1}\boldsymbol{\Phi_{iw}}$, $i = 1, 2$.

Our sub-optimal closed-form solution for the BC under the average power constraint (2) is not optimal, since instead of doing an exhaustive search over all $\mathbf{S} \succeq \underline{0}$ with $\mathrm{Tr}(\mathbf{S}) = P_t$ as indicated in (12), we will only consider specific $\mathbf{S}$ matrices of the form given for $\mathbf{S_w}$ in (23) with diagonal $\mathbf{P}$. Since $R_i^*(\mathbf{P}_i)$ is only a function of $\mathbf{P}_i$, $R_1^*(\mathbf{P}_1)$ and $R_2^*(\mathbf{P}_2)$ can be optimized separately for any power fraction α $(0 \leq \alpha \leq 1)$ under the constraints $\mathrm{Tr}(\mathbf{A}_1\mathbf{P}_1) = \alpha P_t$ and $\mathrm{Tr}(\mathbf{A}_2\mathbf{P}_2) = (1 - \alpha)P_t$, respectively.

Theorem 4. For any α, $0 \leq \alpha \leq 1$, the diagonal elements of the optimal $\mathbf{P}_1^*$ and $\mathbf{P}_2^*$ are given by

$$
p_{1i}^* = \max\left(0, \frac{-(\sigma_{1\rho i} + \sigma_{2\rho i}) + \sqrt{(\sigma_{1\rho i} - \sigma_{2\rho i})^2 + 4(\sigma_{1\rho i} - \sigma_{2\rho i})\sigma_{1\rho i}\sigma_{2\rho i}/(\mu_1 a_{1i})}}{2\,\sigma_{1\rho i}\sigma_{2\rho i}}\right)
\tag{32}
$$

$$
p_{2i}^* = \max\left(0, \frac{-(\sigma_{1\bar{\rho} i} + \sigma_{2\bar{\rho} i}) + \sqrt{(\sigma_{2\bar{\rho} i} - \sigma_{1\bar{\rho} i})^2 + 4(\sigma_{2\bar{\rho} i} - \sigma_{1\bar{\rho} i})\sigma_{2\bar{\rho} i}\sigma_{1\bar{\rho} i}/(\mu_2 a_{2i})}}{2\,\sigma_{1\bar{\rho} i}\sigma_{2\bar{\rho} i}}\right),
\tag{33}
$$

where $\sigma_{1\rho i}$, $\sigma_{2\rho i}$, and a_{1i} are the ith diagonal elements of $\boldsymbol{\Sigma_{1\rho}}$, $\boldsymbol{\Sigma_{2\rho}}$, and $\mathbf{A}_1$, respectively, where $0 \leq i \leq \rho$. Also $\sigma_{1\bar{\rho} i}$, $\sigma_{2\bar{\rho} i}$, and a_{2i} are the ith diagonal elements of $\boldsymbol{\Sigma_{1\bar{\rho}}}$, $\boldsymbol{\Sigma_{2\bar{\rho}}}$, and $\mathbf{A}_2$, respectively, where $0 \leq i \leq (n_t - \rho)$. The Lagrange parameters $\mu_1 > 0$ and $\mu_2 > 0$ are chosen to satisfy the average power constraints $\mathrm{Tr}(\mathbf{A}_1\mathbf{P}_1) = \alpha P_t$ and $\mathrm{Tr}(\mathbf{A}_2\mathbf{P}_2) = (1 - \alpha)P_t$, respectively.

Proof: We want to optimize diagonal matrices $\mathbf{P}_1$ and $\mathbf{P}_2$ so that the secrecy rates $R_1^*(\mathbf{P}_1)$ and $R_2^*(\mathbf{P}_2)$, given by (29), are maximized for a given α, $0 \leq \alpha \leq 1$. Since $R_i^*(\mathbf{P}_i)$ only depends on $\mathbf{P}_i$, the two terms in (29) can be maximized independently. We show the result for $i = 1$; the procedure for $i = 2$ is identical. From (29), the Lagrangian associated with

$\max_{\text{Tr}(\mathbf{A}_1\mathbf{P}_1)=\alpha P_t} R_1^*(\mathbf{P}_1)$ is

$$\mathcal{L} = \log|\mathbf{I} + \mathbf{\Sigma}_{1\rho}\mathbf{P}_1| - |\mathbf{I} + \mathbf{\Sigma}_{2\rho}\mathbf{P}_1| - \mu_1\text{Tr}(\mathbf{A}_1\mathbf{P}_1)$$

$$= \sum_i [\log(1 + \sigma_{1\rho i}p_{1i}) - \log(1 + \sigma_{2\rho i}p_{1i})] - \mu_1\sum_i a_{1i}\,p_{1i}\,, \tag{34}$$

where $\mu_1 > 0$ is the Lagrange multiplier. Since $\mathbf{\Sigma}_{1\rho} \succ \mathbf{\Sigma}_{2\rho}$, Eq. (34) represents a convex optimization problem. The optimal $\mathbf{P}_1^*$ with diagonal elements given by (32) is simply obtained by applying the KKT conditions to (34). ∎

Corollary 1. For any α, $0 \le \alpha \le 1$, let $R_1^*(\alpha)$ and $R_2^*(\alpha)$ represent the corner points given by (29) for the optimal $\mathbf{P}_1^*$ and $\mathbf{P}_2^*$, given by (32) and (33). The achievable secrecy rate region of the above approach under the average power constraint (2) is the convex hull of all obtained corner points and is given by

$$\mathcal{R}_s(\mathbf{H}, \mathbf{G}, P_t) = \bigcup_{0 \le \alpha \le 1} (R_1^*(\alpha)\,, R_2^*(\alpha))\,. \tag{35}$$

It is interesting to note that, unlike the conventional broadcast channel without secrecy constraints where uniform power allocation is optimal in maximizing the sum-rate in the high SNR regime [16], the high SNR power allocation for the BC with confidential messages is a special form of waterfilling as described in the following lemma.

Lemma 4. For high SNR ($P_t \to \infty$), the asymptotic optimal power allocations given by (32)-(33) are

$$p_{1i}^* = \sqrt{\frac{1}{\mu_1 a_{1i}}\left(\frac{1}{\sigma_{2\rho i}} - \frac{1}{\sigma_{1\rho i}}\right)} \tag{36}$$

$$p_{2i}^* = \sqrt{\frac{1}{\mu_2 a_{2i}}\left(\frac{1}{\sigma_{1\bar{\rho} i}} - \frac{1}{\sigma_{2\bar{\rho} i}}\right)}\,. \tag{37}$$

Proof: To show (36) we note that $\mu_1 \to 0$ when $P_t \to \infty$. Thus (32) can be written as

$$p_{1i}^* = \frac{\sqrt{4(\sigma_{1\rho i} - \sigma_{2\rho i})\sigma_{1\rho i}\sigma_{2\rho i}/(\mu_1 a_{1i})}}{2\,\sigma_{1\rho i}\sigma_{2\rho i}} = \sqrt{\frac{1}{\mu_1 a_{1i}}\frac{\sigma_{1\rho i} - \sigma_{2\rho i}}{\sigma_{1\rho i}\sigma_{2\rho i}}} = \sqrt{\frac{1}{\mu_1 a_{1i}}\left(\frac{1}{\sigma_{2\rho i}} - \frac{1}{\sigma_{1\rho i}}\right)}.$$

(37) is proved similarly. ∎

It is also worth noting that the solution in (32)-(33) approaches the standard point-to-point MIMO waterfilling solution when one of the channels is dominant. For example, let $\mathbf{G} \to 0$. We will show that the optimal input covariance simplifies to the waterfilling solution for $\mathbf{H}$, given by $\mathbf{Q}_P^* = \mathbf{\Phi}_H \left(\frac{1}{\mu}\mathbf{I} - \mathbf{\Lambda}_H^{-1}\right)^+ \mathbf{\Phi}_H^H$, where unitary $\mathbf{\Phi}_H$ and diagonal $\mathbf{\Lambda}_H$ are obtained from the eigenvalue decomposition $\mathbf{H}^H\mathbf{H} = \mathbf{\Phi}_H\mathbf{\Lambda}_H\mathbf{\Phi}_H^H$. The capacity of the point-to-point MIMO Gaussian link is

$$C = \log\left|\mathbf{I} + \mathbf{H}\,\mathbf{Q}_P^*\mathbf{H}^H\right| = \log\left|\mathbf{I} + \mathbf{\Lambda}_H\left(\frac{1}{\mu}\mathbf{I} - \mathbf{\Lambda}_H^{-1}\right)^+\right|. \tag{38}$$

When $\mathbf{G} \to \underline{0}$, we note from (20) and (21) that $\mathbf{\Phi_w} \to \mathbf{\Phi}_H$, $\mathbf{\Sigma}_1 = \mathbf{\Sigma}_{1\rho} \to \mathbf{I}$, $\mathbf{\Sigma}_2 \to \underline{0}$, and $\mathbf{P}_1 = \mathbf{P}$. Consequently, $R_2^* \to 0$ and $R_1^* \to \log|\mathbf{I} + \mathbf{P}^*|$, where $\mathbf{P}^*$ is a diagonal matrix with diagonal elements given by (32). The average power constraint in (31) becomes $\mathrm{Tr}\,(\mathbf{AP}) = P_t$, where $\mathbf{A} = \mathbf{\Phi_w}^H(\mathbf{H}^H\mathbf{H} + \mathbf{G}^H\mathbf{G})^{-1}\mathbf{\Phi_w} \to \mathbf{\Lambda}_H^{-1}$ when $\mathbf{G} \to \underline{0}$. Thus the ith diagonal element of $\mathbf{A}$ converges to the ith diagonal element of $\mathbf{\Lambda}_H^{-1}$. Starting from (32) and applying L'Hôpital's rule, when $\mathbf{G} \to 0$ and hence $\sigma_{2\rho i} \to 0$, we have $\mathbf{P}^* \to \mathbf{\Lambda}_H(\frac{1}{\mu}\mathbf{I} - \mathbf{\Lambda}_H^{-1})^+$, and consequently,

$$\lim_{\mathbf{G}\to\underline{0}} R_1^* = \log|\mathbf{I} + \mathbf{P}^*| = \log\left|\mathbf{I} + \mathbf{\Lambda}_H\left(\frac{1}{\mu}\mathbf{I} - \mathbf{\Lambda}_H^{-1}\right)^+\right|.$$

C. Alternative Approach for the MISO Case

Here we focus on the BC in (1) for the MISO case under an average power constraint, where both receivers have a single antenna and the transmitter has $n_t \geq 2$ antennas:

$$y_1 = \mathbf{h}^H\mathbf{x} + z_1$$

$$y_2 = \mathbf{g}^H\mathbf{x} + z_2 \,,$$

where the channels are represented by the $n_t \times 1$ vectors $\mathbf{h}$ and $\mathbf{g}$. The MISO case is the only BC scenario whose secrecy capacity region under the *average* power constraint (2) is characterized in closed-form. In particular, it was shown in [3] that

$$\mathcal{C}_s(\mathbf{h}, \mathbf{g}, P_t) = \bigcup_{0 \leq \alpha \leq 1} (C_1(\alpha), C_2(\alpha)) \tag{39}$$

where $(C_1(\alpha), C_2(\alpha))$ is the secrecy rate pair on the Pareto boundary of the secrecy capacity region for the power fraction α, $0 \leq \alpha \leq 1$, where power αP_t is allocated to receiver 1's message and $(1 - \alpha)P_t$ is allocated to receiver 2's message. Furthermore, we have [3]

$$
\begin{aligned}
C_1(\alpha) &= \log \gamma_1(\alpha) \\
C_2(\alpha) &= \log \gamma_2(\alpha) \, ,
\end{aligned}
\tag{40}
$$

where

$$
\gamma_1(\alpha) = \frac{1 + \alpha P_t \, \mathbf{e}_1^H \mathbf{h}\mathbf{h}^H \mathbf{e}_1}{1 + \alpha P_t \, \mathbf{e}_1^H \mathbf{g}\mathbf{g}^H \mathbf{e}_1} \, ,
$$

$\mathbf{e}_1$ is the unit length principal generalized eigenvector of $(\mathbf{I} + P_t \mathbf{h}\mathbf{h}^H \, ; \, \mathbf{I} + P_t \mathbf{g}\mathbf{g}^H)$, $\gamma_2(\alpha)$ is the largest generalized eigenvalue of

$$
(\mathbf{I} + \frac{(1 - \alpha)P_t}{1 + \alpha P_t |\mathbf{e}_1^H \mathbf{g}|} \mathbf{g}\mathbf{g}^H \, ; \, \mathbf{I} + \frac{(1 - \alpha)P_t}{1 + \alpha P_t |\mathbf{e}_1^H \mathbf{h}|} \mathbf{h}\mathbf{h}^H) \, ,
$$

and $\mathbf{e}_2$ denotes the unit length generalized eigenvector corresponding to $\gamma_2(\alpha)$. Note that the achievablity of (40) is still based on S-DPC.

While we could have just used the results of Section IV-B for the MISO case, we will see that the advantage of considering a different approach here is that we obtain a more succinct expression for the achievable secrecy rate region for linear precoding, and we are able to quantify the loss in secrecy rate incurred by linear precoding under the average power constraint compared with $(C_1(\alpha), C_2(\alpha))$. This was not possible in the MIMO case.

Referring to (6), it was shown in [3] that for the secrecy rate pair given by (40), U_1 and U_2 have covariance matrices $\alpha P_t \mathbf{e}_1 \mathbf{e}_1^H$ and $(1 - \alpha)P_t \mathbf{e}_2 \mathbf{e}_2^H$, respectively. Thus, the specific input covariance matrix that attains (40) is given by

$$
\mathbf{S}_Q = \alpha P_t \mathbf{e}_1 \mathbf{e}_1^H + (1 - \alpha)P_t \mathbf{e}_2 \mathbf{e}_2^H \, ,
\tag{41}
$$

where $\mathrm{Tr}(\mathbf{S}_Q) = P_t$ and $\mathrm{rank}(\mathbf{S}_Q) = 2$. Equivalently, one can say that under the *matrix* power constraint $\mathbf{S}_Q$, the corner point of the corresponding rectangular secrecy capacity region is given by (40). The union of these corner points constructs the Pareto boundary of the secrecy capacity region under the *average* power constraint, where any point on the boundary is given by (40)

for a different α and is achieved under the matrix power constraint $\mathbf{S}_Q$ given by (41).

Using the above fact, we now present a different linear precoding scheme as an alternative to Corollary 1 for the MISO BC under the average power constraint (2).

Corollary 2. Using the linear precoding scheme proposed in Theorem 3 for the MISO BC under an average power constraint, the following secrecy rate region is achievable:

$$\mathcal{R}_s(\mathbf{h}, \mathbf{g}, P_t) = \bigcup_{0 \le \alpha \le 1} (R_1(\alpha), R_2(\alpha)) \ ,$$

where

$$
\begin{aligned}
R_1(\alpha) &= \max(C_1(\alpha) - \log\left(1 + (\mathbf{c}_2^H \mathbf{P}_{\mathbf{c}_1}^\perp \mathbf{c}_2)^{-2} \, |\mathbf{c}_1^H \mathbf{P}_{\mathbf{c}_2}^\perp \mathbf{P}_{\mathbf{c}_1}^\perp \mathbf{c}_2|^2\right), 0) \\
R_2(\alpha) &= \max(C_2(\alpha) - \log\left(1 + (\mathbf{c}_2^H \mathbf{P}_{\mathbf{c}_1}^\perp \mathbf{c}_2)^{-2} \, |\mathbf{c}_1^H \mathbf{P}_{\mathbf{c}_2}^\perp \mathbf{P}_{\mathbf{c}_1}^\perp \mathbf{c}_2|^2\right), 0) \ ,
\end{aligned}
\tag{42}
$$

$C_1(\alpha)$ and $C_2(\alpha)$ are given by (40),

$$\mathbf{c}_1 = \frac{1}{\sqrt{\mathbf{e}_1^H(\mathbf{S}_Q^{-1} + \mathbf{g}\mathbf{g}^H)\mathbf{e}_1}} \, \mathbf{S}_Q^{-\frac{1}{2}} \mathbf{e}_1 \tag{43}$$

$$\mathbf{c}_2 = \frac{1}{\sqrt{\mathbf{f}_1^H(\mathbf{S}_Q^{-1} + \mathbf{g}\mathbf{g}^H)\mathbf{f}_1}} \, \mathbf{S}_Q^{-\frac{1}{2}} \mathbf{f}_1 \ , \tag{44}$$

and where $\mathbf{f}_1$ is the unit length principal generalized eigenvector of $(\mathbf{I} + P_t \mathbf{g}\mathbf{g}^H \ ; \ \mathbf{I} + P_t \mathbf{h}\mathbf{h}^H)$.

Proof: From Remark 2, and by noting that for any MISO BC, $\mathbf{h}\mathbf{h}^H + \mathbf{g}\mathbf{g}^H$ has at most 2 non-zero eigenvalues, any MISO BC can be modeled with a scenario involving just two transmit antennas. Thus, without loss of generality, we assume that $n_t = 2$. From Theorem 3, we only need to characterize $\mathbf{c}_1$ and $\mathbf{c}_2$, where $\mathbf{c}_1$ ($\mathbf{c}_2$) is the generalized eigenvector of the pencil

$$\left(\mathbf{S}_Q^{\frac{1}{2}} \mathbf{h}\mathbf{h}^H \mathbf{S}_Q^{\frac{1}{2}} + \mathbf{I} \ , \ \mathbf{S}_Q^{\frac{1}{2}} \mathbf{g}\mathbf{g}^H \mathbf{S}_Q^{\frac{1}{2}} + \mathbf{I}\right) \tag{45}$$

corresponding to the generalized eigenvalue larger (less) than 1, λ_1 (λ_2).

From (6) and (7), the covariance matrix of U_1 can be rewritten as

$$\mathbf{K}_t^* = \mathbf{S}_Q^{\frac{1}{2}} \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix} \begin{bmatrix} (\mathbf{c}_1^H \mathbf{c}_1)^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{c}_1 & \mathbf{c}_2 \end{bmatrix}^H \mathbf{S}_Q^{\frac{1}{2}} = \frac{1}{\mathbf{c}_1^H \mathbf{c}_1} \mathbf{S}_Q^{\frac{1}{2}} \mathbf{c}_1 \, \mathbf{c}_1^H \mathbf{S}_Q^{\frac{1}{2}} \ . \tag{46}$$
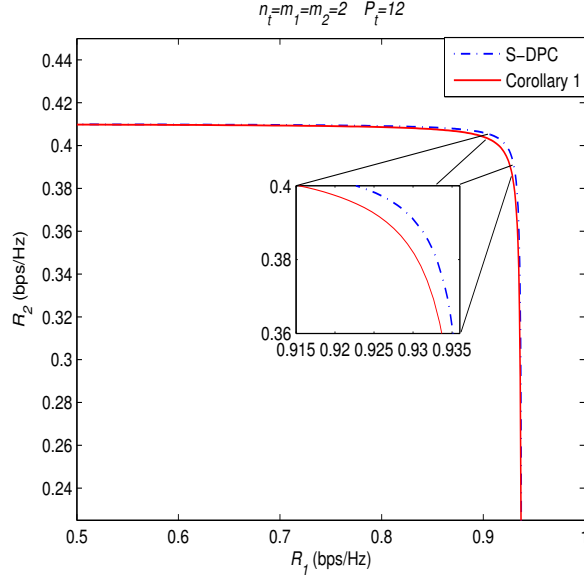
Fig. 1. Secrecy capacity region of S-DPC together with secrecy rate region for linear precoding with $P_t = 12$, $n_t = m_1 = m_2 = 2$.

Comparing (46) with the covariance matrix of U_1 reported in [3], we have $\alpha P_t \mathbf{e}_1 \mathbf{e}_1^H = \frac{1}{|\mathbf{c}_1|^2} \mathbf{S}_Q^{\frac{1}{2}} \mathbf{c}_1 \, \mathbf{c}_1^H \mathbf{S}_Q^{\frac{1}{2}}$. This results in[4]

$$\frac{\mathbf{c}_1}{\|\mathbf{c}_1\|} = \sqrt{\alpha P_t} \, \mathbf{S}_Q^{-\frac{1}{2}} \mathbf{e}_1 \ . \tag{47}$$

On the other hand, from the definition of $\mathbf{c}_1$ and $\mathbf{c}_2$ (see (9)-(10) for example), we have

$$[\mathbf{c}_1 \ \mathbf{c}_2]^H \left[\mathbf{S}_Q^{\frac{1}{2}} \mathbf{h}\mathbf{h}^H \mathbf{S}_Q^{\frac{1}{2}} + \mathbf{I} \right] [\mathbf{c}_1 \ \mathbf{c}_2] = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} = \begin{bmatrix} \gamma_1(\alpha) & 0 \\ 0 & \gamma_2^{-1}(\alpha) \end{bmatrix}$$
$$[\mathbf{c}_1 \ \mathbf{c}_2]^H \left[\mathbf{S}_Q^{\frac{1}{2}} \mathbf{g}\mathbf{g}^H \mathbf{S}_Q^{\frac{1}{2}} + \mathbf{I} \right] [\mathbf{c}_1 \ \mathbf{c}_2] = \mathbf{I} \tag{48}$$

where $\gamma_1(\alpha)$ and $\gamma_2(\alpha)$ are defined after (40), and the fact that $\lambda_1 = \gamma_1(\alpha)$ and $\lambda_2 = \frac{1}{\gamma_2(\alpha)}$ comes from the argument after (41) and by comparing (11) and (40). Substituting (47) in (48), after some simple calculations, $\mathbf{c}_1$ can be explicitly written as in (43). Recalling that $\mathbf{c}_1$ is the principal generalized eigenvector of (45) and $\mathbf{c}_2$, which corresponds to the smallest generalized

---

[4]Note that multiplication by a factor $\exp(j\theta)$ is required for a precise equaltiy, but since this term disappears in the final result, we simply ignore it.

eigenvalue of the pencil (45), is the principal generalized eigenvector of the pencil

$$\left(\mathbf{S}_Q^{\frac{1}{2}} \mathbf{g} \mathbf{g}^H \mathbf{S}_Q^{\frac{1}{2}} + \mathbf{I} \,, \, \mathbf{S}_Q^{\frac{1}{2}} \mathbf{h} \mathbf{h}^H \mathbf{S}_Q^{\frac{1}{2}} + \mathbf{I} \right) \,,$$

we obtain (44). The proof is completed by using (43) and (44) in (19). ∎
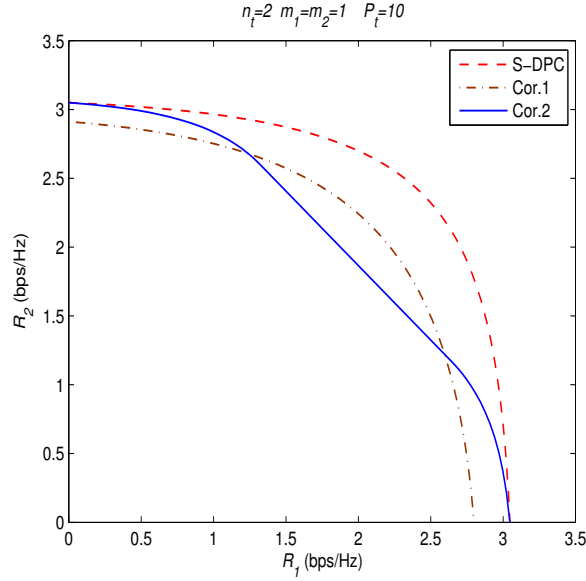
V. NUMERICAL RESULTS

Fig. 2. Secrecy capacity region of S-DPC together with secrecy rate region for linear precodings in Cor. 1 and Cor.2, with $P_t = 10$, $n_t = 2$, $m_1 = m_2 = 1$.

In this section, we provide numerical examples to illustrate the achievable secrecy rate region of the MIMO Gaussian BC under the average power constraint (2). In the first example, we have $P_t = 12$, $\mathbf{H} = [0.3\ 2.5; 2.2\ 1.8]$ and $\mathbf{G} = [1.3\ 1.2; 1.5\ 3.9]$, which is identical to the case studied in [4, Fig. 3 (d)]. Fig. 1 compares the achievable secrecy rate region of the proposed linear precoding scheme in Section IV-A with the secrecy capacity region obtained by the optimal S-DPC approach together with an exhaustive search over suitable matrix constraints, as described in Section II. We see that in this example, the performance of the proposed linear precoding approach is essentially identical to that of the optimal S-DPC scheme.

In the next example, we study the MISO BC for $P_t = 10$. Fig. 2 shows the average secrecy rate regions for S-DPC and the suboptimal linear precoding algorithms described in Corollary 1

and 2. This plot is based on an average of over 30000 channel realizations, where the channel coefficients were generated as independent $\mathcal{CN}(0,1)$ random variables. We see that Corollary 2 provides near optimal performance when $\alpha \to 0$ or $\alpha \to 1$, while Corollary 1 is better for in-between values of α. The degradation of using linear precoding with Corollary 1 is never above 15% for any α.

VI. CONCLUSIONS

We have shown that for a two-user Gaussian BC with an arbitrary number of antennas at each node, when the channel input is under the matrix power constraint, linear precoding is optimal and achieves the secrecy capacity region attained by the optimal S-DPC approach if the matrix constraint satisfies a specific condition. We characterized the form of the linear precoding that achieves the secrecy capacity region in such cases, and we quantified the maximum loss in secrecy rate that occurs if the matrix power constraint does not satisfy the given condition. Based on these observations, we then formulated a sub-optimal approach for the general MIMO scenario based on linear precoding for the case of an average power constraint, for which no known characterization of the secrecy capacity region exists. We also studied the MISO case in detail. Numerical results indicate that the proposed linear precoding approaches yield secrecy rate regions that are close to the secrecy capacity achieved by S-DPC.

APPENDIX A
PROOF OF EQ. (15)

From (14), we have

$$
\begin{aligned}
R_1 &= I(V_1;\, \mathbf{H}(V_1 + V_2) + Z_1) - I(V_1;\, \mathbf{G}V_1 + Z_2) \\
&= \log\left|\mathbf{I} + \mathbf{S}\mathbf{H}^H\mathbf{H}\right| - \log\left|\mathbf{I} + (\mathbf{S} - \mathbf{K}_t^*)\mathbf{H}^H\mathbf{H}\right| - \log\left|\mathbf{I} + \mathbf{K}_t^*\mathbf{G}^H\mathbf{G}\right| \ . \quad (49)
\end{aligned}
$$

The covariance $\mathbf{K}_t^*$, given by (7), can be rewritten as

$$
\mathbf{K}_t^* = \mathbf{S}^{\frac{1}{2}} \begin{bmatrix} \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix} \begin{bmatrix} (\mathbf{C}_1^H\mathbf{C}_1)^{-1} & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \begin{bmatrix} \mathbf{C}_1^H \\ \mathbf{C}_2^H \end{bmatrix} \mathbf{S}^{\frac{1}{2}}
$$

$$= \mathbf{S}^{\frac{1}{2}} \mathbf{C}_1 (\mathbf{C}_1^H \mathbf{C}_1)^{-1} \mathbf{C}_1^H \mathbf{S}^{\frac{1}{2}} = \mathbf{S}^{\frac{1}{2}} \mathbf{P}_{\mathbf{C}_1} \mathbf{S}^{\frac{1}{2}} \,, \tag{50}$$

where $\mathbf{P}_{\mathbf{C}_1} = \mathbf{C}_1 (\mathbf{C}_1^H \mathbf{C}_1)^{-1} \mathbf{C}_1^H$ is the projection matrix onto the column space of $\mathbf{C}_1$. Moreover, let $\mathbf{P}_{\mathbf{C}_1}^{\perp} = \mathbf{I} - \mathbf{P}_{\mathbf{C}_1}$ be the projection onto the space orthogonal to $\mathbf{C}_1$. Consequently, we have

$$\mathbf{S} - \mathbf{K}_t^* = \mathbf{S} - \mathbf{S}^{\frac{1}{2}} \mathbf{P}_{\mathbf{C}_1} \mathbf{S}^{\frac{1}{2}}$$

$$= \mathbf{S}^{\frac{1}{2}} \mathbf{P}_{\mathbf{C}_1}^{\perp} \mathbf{S}^{\frac{1}{2}} = \mathbf{S}^{\frac{1}{2}} \mathbf{P}_{\mathbf{C}_2} \mathbf{S}^{\frac{1}{2}} \tag{51}$$

$$= \mathbf{S}^{\frac{1}{2}} \mathbf{C} \begin{bmatrix} \underline{0} & \underline{0} \\ \underline{0} & (\mathbf{C}_2^H \mathbf{C}_2)^{-1} \end{bmatrix} \mathbf{C}^H \mathbf{S}^{\frac{1}{2}} \,, \tag{52}$$

where in (51), $\mathbf{P}_{\mathbf{C}_1}^{\perp} = \mathbf{P}_{\mathbf{C}_2}$ comes from the fact that $\mathrm{span}\{\mathbf{C}_1\} \perp \mathrm{span}\{\mathbf{C}_2\}$, and $\mathbf{C} = [\mathbf{C}_1 \quad \mathbf{C}_2]$ is full-rank.

Following the same steps as in the proof of [13, Lemma 2] or [4, App. B], we can convert the case when $\mathbf{S} \succeq \underline{0}$, $|\mathbf{S}| = 0$, to the case where $\mathbf{S} \succ \underline{0}$ with the same secrecy capacity region. From (9) and (10) we have

$$\mathbf{H}^H \mathbf{H} = \mathbf{S}^{-1/2} \left[\mathbf{C}^{-H} \begin{bmatrix} \mathbf{\Lambda}_1 & 0 \\ 0 & \mathbf{\Lambda}_2 \end{bmatrix} \mathbf{C}^{-1} - \mathbf{I} \right] \mathbf{S}^{-1/2} \tag{53}$$

$$\mathbf{G}^H \mathbf{G} = \mathbf{S}^{-1/2} \left[\mathbf{C}^{-H} \mathbf{C}^{-1} - \mathbf{I} \right] \mathbf{S}^{-1/2} \,.$$

Using (52) and (53), we have:

$$\left| \mathbf{I} + (\mathbf{S} - \mathbf{K}_t^*) \mathbf{H}^H \mathbf{H} \right| = \left| \mathbf{I} + \mathbf{S}^{\frac{1}{2}} \mathbf{C} \begin{bmatrix} \underline{0} & \underline{0} \\ \underline{0} & (\mathbf{C}_2^H \mathbf{C}_2)^{-1} \end{bmatrix} \mathbf{C}^H \cdot \left[\mathbf{C}^{-H} \begin{bmatrix} \mathbf{\Lambda}_1 & 0 \\ 0 & \mathbf{\Lambda}_2 \end{bmatrix} \mathbf{C}^{-1} - \mathbf{I} \right] \mathbf{S}^{-1/2} \right|$$

$$= \left| \mathbf{I} + \begin{bmatrix} \underline{0} & \underline{0} \\ \underline{0} & (\mathbf{C}_2^H \mathbf{C}_2)^{-1} \end{bmatrix} \cdot \left[\begin{bmatrix} \mathbf{\Lambda}_1 & 0 \\ 0 & \mathbf{\Lambda}_2 \end{bmatrix} - \mathbf{C}^H \mathbf{C} \right] \right| \tag{54}$$

$$= \left| \begin{bmatrix} \mathbf{I} & \underline{0} \\ \underline{0} & (\mathbf{C}_2^H \mathbf{C}_2)^{-1} \mathbf{\Lambda}_2 \end{bmatrix} \right| \tag{55}$$

$$= \left| (\mathbf{C}_2^H \mathbf{C}_2)^{-1} \mathbf{\Lambda}_2 \right| = \left| (\mathbf{C}_2^H \mathbf{C}_2)^{-1} \right| \cdot |\mathbf{\Lambda}_2| \,, \tag{56}$$

where (54) comes from the fact that $|\mathbf{I} + \mathbf{A}\mathbf{B}| = |\mathbf{I} + \mathbf{B}\mathbf{A}|$. Finally, (55) holds since $\mathbf{C}_1^H \mathbf{C}_2 = \underline{0}$

and $\mathbf{C}^H\mathbf{C}$ is block diagonal.

Similarly, one can show that

$$\left|\mathbf{I} + \mathbf{K}_t^*\mathbf{G}^H\mathbf{G}\right| = \left|(\mathbf{C}_1^H\mathbf{C}_1)^{-1}\right| \tag{57}$$

and

$$\left|\mathbf{I} + \mathbf{S}\mathbf{H}^H\mathbf{H}\right| = \left|\mathbf{C}^{-H}\begin{bmatrix}\boldsymbol{\Lambda}_1 & 0 \\ 0 & \boldsymbol{\Lambda}_2\end{bmatrix}\mathbf{C}^{-1}\right| = \left|(\mathbf{C}^H\mathbf{C})^{-1}\right| \cdot |\boldsymbol{\Lambda}_1| \cdot |\boldsymbol{\Lambda}_2|$$

$$= \left|(\mathbf{C}_1^H\mathbf{C}_1)^{-1}\right| \cdot \left|(\mathbf{C}_2^H\mathbf{C}_2)^{-1}\right| \cdot |\boldsymbol{\Lambda}_1| \cdot |\boldsymbol{\Lambda}_2| \ . \tag{58}$$

Substituting (56), (57) and (58) in (49), we have $R_1 = \log|\boldsymbol{\Lambda}_1| = R_1^*$, and this completes the proof.

Appendix B

Proof of Lemma 1

From (9)-(10), we know that $\mathrm{rank}(\mathbf{C}_1) = b$, where b represents number of generalized eigenvalues of the pencil (8) that are greater than 1. From (9)-(10), we have

$$\mathbf{C}_1^H\left[\mathbf{S}^{\frac{1}{2}}\mathbf{H}^H\mathbf{H}\mathbf{S}^{\frac{1}{2}} + \mathbf{I}\right]\mathbf{C}_1 = \boldsymbol{\Lambda}_1 \tag{59}$$

$$\mathbf{C}_1^H\left[\mathbf{S}^{\frac{1}{2}}\mathbf{G}^H\mathbf{G}\mathbf{S}^{\frac{1}{2}} + \mathbf{I}\right]\mathbf{C}_1 = \mathbf{I} \ . \tag{60}$$

Subtracting (59) from (60), a straightforward computation yields

$$\mathbf{C}_1^H\mathbf{S}^{\frac{1}{2}}\left[\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G}\right]\mathbf{S}^{\frac{1}{2}}\mathbf{C}_1 = \boldsymbol{\Lambda}_1 - \mathbf{I} \succ \underline{0} \ . \tag{61}$$

From (61), we have $\mathbf{C}_1^H\mathbf{S}^{\frac{1}{2}}\left[\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G}\right]\mathbf{S}^{\frac{1}{2}}\mathbf{C}_1 \succ \underline{0}$, from which it follows that $\mathrm{rank}(\mathbf{C}_1) = b \leq m$. Similarly one can show that $\mathrm{rank}(\mathbf{C}_2') \leq m'$, where $\mathbf{C}_2'$ corresponds to the generalized eigenvalues of the pencil (8) which are less than 1, and m' represents number of negative eigenvalues of $\mathbf{H}^H\mathbf{H} - \mathbf{G}^H\mathbf{G}$.

APPENDIX C

PROOF OF THEOREM 2

We want to characterize the matrices $\mathbf{S} \succeq \underline{0}$ for which

$$\left(\mathbf{S}^{\frac{1}{2}} \mathbf{H}^H \mathbf{H} \mathbf{S}^{\frac{1}{2}} + \mathbf{I} \,,\; \mathbf{S}^{\frac{1}{2}} \mathbf{G}^H \mathbf{G} \mathbf{S}^{\frac{1}{2}} + \mathbf{I} \right) \tag{62}$$

has generalized eigenvectors with orthogonal $\mathbf{C}_1$ and $\mathbf{C}_2$. For any positive semidefinite matrix $\mathbf{S} \in \mathbb{C}^{n_t \times n_t}$, there exists a matrix $\mathbf{T} \in \mathbb{C}^{n_t \times n_t}$ such that $\mathbf{S} = \mathbf{T}\mathbf{T}^H$ [12]. More precisely, $\mathbf{T} = \mathbf{S}^{\frac{1}{2}} \boldsymbol{\Psi}$, where $\boldsymbol{\Psi}$ can be any $n_t \times n_t$ unitary matrix; thus $\mathbf{T}$ is not unique.

Remark 4. Let the invertible matrix $\overline{\mathbf{C}}$ and the diagonal matrix $\overline{\boldsymbol{\Lambda}}$ respectively represent the generalized eigenvectors and eigenvalues of

$$\left(\mathbf{T}^H \mathbf{H}^H \mathbf{H} \mathbf{T} + \mathbf{I} \,,\; \mathbf{T}^H \mathbf{G}^H \mathbf{G} \mathbf{T} + \mathbf{I} \right) \,, \tag{63}$$

so that

$$\begin{aligned}
\overline{\mathbf{C}}^H \left[\mathbf{T}^H \mathbf{H}^H \mathbf{H} \mathbf{T} + \mathbf{I} \right] \overline{\mathbf{C}} &= \overline{\boldsymbol{\Lambda}} \\
\overline{\mathbf{C}}^H \left[\mathbf{T}^H \mathbf{G}^H \mathbf{G} \mathbf{T} + \mathbf{I} \right] \overline{\mathbf{C}} &= \mathbf{I} \,,
\end{aligned} \tag{64}$$

where $\mathbf{T} = \mathbf{S}^{\frac{1}{2}} \boldsymbol{\Psi}$ for a given unitary matrix $\boldsymbol{\Psi}$. By comparing (9) and (64), one can confirm that $\boldsymbol{\Psi}\overline{\mathbf{C}} = \mathbf{C}$ and $\overline{\boldsymbol{\Lambda}} = \boldsymbol{\Lambda}$, where $\mathbf{C}$ and $\boldsymbol{\Lambda}$ are respectively the generalized eigenvectors and eigenvalues of (62), as given by (9).

Also note that, for any unitary $\boldsymbol{\Psi}$, $\overline{\mathbf{C}}^H \overline{\mathbf{C}} = \mathbf{C}^H \mathbf{C}$. Thus, finding a $\mathbf{S} \succeq \underline{0}$ such that (62) has orthogonal $\mathbf{C}_1$ and $\mathbf{C}_2$ (block diagonal $\mathbf{C}^H \mathbf{C}$) is equivalent to finding a $\mathbf{T}$, $\mathbf{S} = \mathbf{T}\mathbf{T}^H$, such that (63) has orthogonal $\overline{\mathbf{C}}_1$ and $\overline{\mathbf{C}}_2$ (block diagonal $\overline{\mathbf{C}}^H \overline{\mathbf{C}}$).

The *if* part of Theorem 2 is easy to show. We want to show that if $\mathbf{S} = \mathbf{T}\mathbf{T}^H$ and $\mathbf{T}$ simultaneously block diagonalizes $\mathbf{H}^H \mathbf{H}$ and $\mathbf{G}^H \mathbf{G}$, as given by (16) such that $\mathbf{K}_{\mathbf{H}1} \succeq \mathbf{K}_{\mathbf{G}1}$ and $\mathbf{K}_{\mathbf{H}2} \preceq \mathbf{K}_{\mathbf{G}2}$, then $\mathbf{C}_1^H \mathbf{C}_2 = \underline{0}$. From the definition of the generalized eigenvalue decomposition,

we have

$$\overline{\mathbf{C}}^H \left[\mathbf{T}^H \mathbf{H}^H \mathbf{H} \mathbf{T} + \mathbf{I} \right] \overline{\mathbf{C}} = \overline{\mathbf{C}}^H \begin{bmatrix} \mathbf{I} + \mathbf{K_{H1}} & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \mathbf{I} + \mathbf{K_{H2}} \end{bmatrix} \overline{\mathbf{C}} = \begin{bmatrix} \mathbf{D}_1 & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \mathbf{D}_2 \end{bmatrix}$$

$$\overline{\mathbf{C}}^H \left[\mathbf{T}^H \mathbf{G}^H \mathbf{G} \mathbf{T} + \mathbf{I} \right] \overline{\mathbf{C}} = \overline{\mathbf{C}}^H \begin{bmatrix} \mathbf{I} + \mathbf{K_{G1}} & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \mathbf{I} + \mathbf{K_{G2}} \end{bmatrix} \overline{\mathbf{C}} = \begin{bmatrix} \mathbf{I} & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \mathbf{I} \end{bmatrix} , \tag{65}$$

from which we have

$$\overline{\mathbf{C}} = \begin{bmatrix} \overline{\mathbf{C}}_1 & \overline{\mathbf{C}}_2 \end{bmatrix} = \begin{bmatrix} \overline{\mathbf{C}}_{11} & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \overline{\mathbf{C}}_{22} \end{bmatrix} ,$$

where the invertible matrix $\overline{\mathbf{C}}_{11}$ and diagonal matrix $\mathbf{D}_1$ are respectively the generalized eigen-vectors and eigenvalues of $(\mathbf{I} + \mathbf{K_{H1}} ; \mathbf{I} + \mathbf{K_{G1}})$. Since $\mathbf{K_{H1}} \succeq \mathbf{K_{G1}}$, then $\mathbf{D}_1 \succeq \mathbf{I}$, which shows that $\overline{\mathbf{C}}_1$ corresponds to generalized eigenvalues that are bigger than or equal to one. We have a similar definition for $\overline{\mathbf{C}}_{22}$ and diagonal matrix $\mathbf{D}_2$, corresponding to $(\mathbf{I} + \mathbf{K_{H2}} ; \mathbf{I} + \mathbf{K_{G2}})$. Finally, since $\overline{\mathbf{C}}^H \overline{\mathbf{C}}$ is block diagonal, then $\mathbf{C}^H \mathbf{C}$, where $\mathbf{C}$ is the generalized eigenvector matrix of (62), is block diagonal as well. This completes the *if* part of the theorem.

In the following, we prove the *only if* part of Theorem 2; *i.e.,* we show that if $\mathbf{S} \succeq \underline{\mathbf{0}}$ results in (62) having orthogonal $\mathbf{C}_1$ and $\mathbf{C}_2$, then there must exist a square matrix $\mathbf{T}$ such that $\mathbf{S} = \mathbf{T}\mathbf{T}^H$ and $\mathbf{T}^H \mathbf{H}^H \mathbf{H} \mathbf{T}$ and $\mathbf{T}^H \mathbf{G}^H \mathbf{G} \mathbf{T}$ are simultaneously block diagonalized as in (16) with $\mathbf{K_{H1}} \succeq \mathbf{K_{G1}}$ and $\mathbf{K_{H2}} \preceq \mathbf{K_{G2}}$.

Let $\mathbf{S}^{\frac{1}{2}} \mathbf{G}^H \mathbf{G} \mathbf{S}^{\frac{1}{2}}$ have the eigenvalue decomposition $\mathbf{\Phi}_B \mathbf{\Sigma}_B \mathbf{\Phi}_B^H$, where $\mathbf{\Phi}_B$ is unitary and $\mathbf{\Sigma}_B$ is a positive semidefinite diagonal matrix. Also let $(\mathbf{I} + \mathbf{\Sigma}_B)^{-\frac{1}{2}} \mathbf{\Phi}_B^H \left(\mathbf{I} + \mathbf{S}^{\frac{1}{2}} \mathbf{H}^H \mathbf{H} \mathbf{S}^{\frac{1}{2}} \right) \mathbf{\Phi}_B (\mathbf{I} + \mathbf{\Sigma}_B)^{-\frac{1}{2}}$ have the eigenvalue decomposition $\mathbf{\Phi}_A \mathbf{\Sigma}_A \mathbf{\Phi}_A^H$, where $\mathbf{\Phi}_A$ is unitary and $\mathbf{\Sigma}_A$ is a positive definite diagonal matrix. One can easily confirm that [12] $\mathbf{C} = \mathbf{\Phi}_B (\mathbf{I} + \mathbf{\Sigma}_B)^{-\frac{1}{2}} \mathbf{\Phi}_A$ and $\mathbf{\Lambda} = \mathbf{\Sigma}_A$, where $\mathbf{C}$ and $\mathbf{\Lambda}$ are respectively the generalized eigenvectors and eigenvalues of (62). Also let $\mathbf{C}$ be ordered such that $\mathbf{C} = [\mathbf{C}_1 \quad \mathbf{C}_2]$, where $\mathbf{C}_1$ corresponds to the generalized eigenvalues bigger than (or equal to) 1. We have

$$\mathbf{C}^H \mathbf{C} = \mathbf{\Phi}_A^H (\mathbf{I} + \mathbf{\Sigma}_B)^{-1} \mathbf{\Phi}_A . \tag{66}$$

From (66), $\mathbf{C}^H \mathbf{C}$ is block diagonal iff the unitary matrix $\mathbf{\Phi}_A$ is block diagonal. Recall-

ing that $\mathbf{\Phi}_A$ is the eigenvector matrix of $(\mathbf{I}+\mathbf{\Sigma}_B)^{-\frac{1}{2}}\mathbf{\Phi}_B^H\left(\mathbf{I}+\mathbf{S}^{\frac{1}{2}}\mathbf{H}^H\mathbf{H}\mathbf{S}^{\frac{1}{2}}\right)\mathbf{\Phi}_B(\mathbf{I}+\mathbf{\Sigma}_B)^{-\frac{1}{2}}$, a block diagonal $\mathbf{\Phi}_A$ leads to $(\mathbf{I}+\mathbf{\Sigma}_B)^{-\frac{1}{2}}\mathbf{\Phi}_B^H\left(\mathbf{I}+\mathbf{S}^{\frac{1}{2}}\mathbf{H}^H\mathbf{H}\mathbf{S}^{\frac{1}{2}}\right)\mathbf{\Phi}_B(\mathbf{I}+\mathbf{\Sigma}_B)^{-\frac{1}{2}}$, and consequently $\mathbf{\Phi}_B^H\mathbf{S}^{\frac{1}{2}}\mathbf{H}^H\mathbf{H}\mathbf{S}^{\frac{1}{2}}\mathbf{\Phi}_B$ must be block diagonal. Thus, if $\mathbf{C}^H\mathbf{C}$ is block diagonal, *i.e.,* $\mathbf{C}_1^H\mathbf{C}_2 = \underline{0}$, there must exist a unitary matrix $\mathbf{\Phi}_B$ such that $\mathbf{\Phi}_B^H\mathbf{S}^{\frac{1}{2}}\mathbf{H}^H\mathbf{H}\mathbf{S}^{\frac{1}{2}}\mathbf{\Phi}_B$ and $\mathbf{\Phi}_B^H\mathbf{S}^{\frac{1}{2}}\mathbf{G}^H\mathbf{G}\mathbf{S}^{\frac{1}{2}}\mathbf{\Phi}_B$ are simultaneously block diagonal.[5] Letting $\mathbf{T} = \mathbf{S}^{\frac{1}{2}}\mathbf{\Phi}_B$ results in (65), for which we must have $\mathbf{K_{H1}} \succeq \mathbf{K_{G1}}$ and $\mathbf{K_{H2}} \preceq \mathbf{K_{G2}}$, otherwise it contradicts the ordering of $\mathbf{C} = [\mathbf{C}_1 \quad \mathbf{C}_2]$. This completes the proof.

APPENDIX D
PROOF OF THEOREM 3

We need to prove that the secrecy rate pair (R_1, R_2) given by (19) is achievable.

Remark 5. By applying the Schur Complement Lemma [12] on

$$\mathbf{C}^H\mathbf{C} = [\mathbf{C}_1 \quad \mathbf{C}_2]^H [\mathbf{C}_1 \quad \mathbf{C}_2] = \begin{bmatrix} \mathbf{C}_1^H\mathbf{C}_1 & \mathbf{C}_1^H\mathbf{C}_2 \\ \mathbf{C}_2^H\mathbf{C}_1 & \mathbf{C}_2^H\mathbf{C}_2 \end{bmatrix}$$

and recalling the fact that $\mathbf{C}$ is full-rank, we have that $\mathbf{C}_2^H\mathbf{C}_2 - \mathbf{C}_2^H\mathbf{C}_1(\mathbf{C}_1^H\mathbf{C}_1)^{-1}\mathbf{C}_1 = \mathbf{C}_2^H\mathbf{P}_{\mathbf{C}_1}^\perp\mathbf{C}_2$ is full rank. Similarly, one can show that $\left(\mathbf{C}_1^H\mathbf{P}_{\mathbf{C}_2}^\perp\mathbf{C}_1\right)^{-1}$ exists. Also, we have $\left|\mathbf{C}^H\mathbf{C}\right| = \left|\mathbf{C}_1^H\mathbf{P}_{\mathbf{C}_2}^\perp\mathbf{C}_1\right| \cdot \left|\mathbf{C}_2^H\mathbf{C}_2\right| = \left|\mathbf{C}_2^H\mathbf{P}_{\mathbf{C}_1}^\perp\mathbf{C}_2\right| \cdot \left|\mathbf{C}_1^H\mathbf{C}_1\right|$.

Define $\widehat{\mathbf{C}} = [\mathbf{P}_{\mathbf{C}_2}^\perp\mathbf{C}_1 \quad \mathbf{C}_2]$, so that

$$\widehat{\mathbf{C}}^H\widehat{\mathbf{C}} = \begin{bmatrix}\mathbf{P}_{\mathbf{C}_2}^\perp\mathbf{C}_1 & \mathbf{C}_2\end{bmatrix}^H \begin{bmatrix}\mathbf{P}_{\mathbf{C}_2}^\perp\mathbf{C}_1 & \mathbf{C}_2\end{bmatrix} = \begin{bmatrix} \mathbf{C}_1^H\mathbf{P}_{\mathbf{C}_2}^\perp\mathbf{C}_1 & \underline{0} \\ \underline{0} & \mathbf{C}_2^H\mathbf{C}_2 \end{bmatrix}.$$

Consequently, we can write

$$\mathbf{P}_{\mathbf{C}_2} = \mathbf{C}_2(\mathbf{C}_2^H\mathbf{C}_2)^{-1}\mathbf{C}_2^H = \widehat{\mathbf{C}} \begin{bmatrix} \underline{0} & 0 \\ 0 & (\mathbf{C}_2^H\mathbf{C}_2)^{-1} \end{bmatrix} \widehat{\mathbf{C}}^H \tag{67}$$

[5]Note that $\mathbf{\Phi}_B^H\mathbf{S}^{\frac{1}{2}}\mathbf{G}^H\mathbf{G}\mathbf{S}^{\frac{1}{2}}\mathbf{\Phi}_B$ is actually diagonal, and hence also block diagonal.

and

$$\mathbf{P}_{\mathbf{C}_2}^{\perp} = \mathbf{I} - \mathbf{P}_{\mathbf{C}_2} = \widehat{\mathbf{C}} \left(\widehat{\mathbf{C}}^H \widehat{\mathbf{C}} \right)^{-1} \widehat{\mathbf{C}}^H - \mathbf{P}_{\mathbf{C}_2}$$

$$= \widehat{\mathbf{C}} \begin{bmatrix} \left(\mathbf{C}_1^H \mathbf{P}_{\mathbf{C}_2}^{\perp} \mathbf{C}_1 \right)^{-1} & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \widehat{\mathbf{C}}^H . \tag{68}$$

In the following we show the achievablity of R_1 in (19). The achievablity of R_2 is obtained in a similar manner. Since V_1 and V_2 in Theorem 3 are independent, from (14) we have

$$R_1 = I(V_1; \mathbf{H}(V_1 + V_2) + Z_1) - I(V_1; \mathbf{G}V_1 + Z_2)$$

$$= \log \left| \mathbf{I} + \mathbf{H} \mathbf{S} \mathbf{H}^H \right| - \log \left| \mathbf{I} + \mathbf{H}(\mathbf{S}^{\frac{1}{2}} \mathbf{P}_{\mathbf{C}_2} \mathbf{S}^{\frac{1}{2}}) \mathbf{H}^H \right| - \log \left| \mathbf{I} + \mathbf{G}(\mathbf{S}^{\frac{1}{2}} \mathbf{P}_{\mathbf{C}_2}^{\perp} \mathbf{S}^{\frac{1}{2}}) \mathbf{G}^H \right|$$

$$= \log \left| \mathbf{I} + \mathbf{S} \mathbf{H}^H \mathbf{H} \right| - \log \left| \mathbf{I} + \mathbf{S}^{\frac{1}{2}} \mathbf{P}_{\mathbf{C}_2} \mathbf{S}^{\frac{1}{2}} \mathbf{H}^H \mathbf{H} \right| - \log \left| \mathbf{I} + \mathbf{S}^{\frac{1}{2}} \mathbf{P}_{\mathbf{C}_2}^{\perp} \mathbf{S}^{\frac{1}{2}} \mathbf{G}^H \mathbf{G} \right| . \tag{69}$$

Recalling (53), we have

$$\left| \mathbf{I} + \mathbf{S} \mathbf{H}^H \mathbf{H} \right| = \left| (\mathbf{C}^H \mathbf{C})^{-1} \right| \cdot |\mathbf{\Lambda}_1| \cdot |\mathbf{\Lambda}_2|$$

$$= \left| (\mathbf{C}_1^H \mathbf{P}_{\mathbf{C}_2}^{\perp} \mathbf{C}_1)^{-1} \right| \cdot \left| (\mathbf{C}_2^H \mathbf{C}_2)^{-1} \right| \cdot |\mathbf{\Lambda}_1| \cdot |\mathbf{\Lambda}_2| , \tag{70}$$

where we used Remark 5 to obtain (70). From (67), we have

$$\log \left| \mathbf{I} + \mathbf{S}^{\frac{1}{2}} \mathbf{P}_{\mathbf{C}_2} \mathbf{S}^{\frac{1}{2}} \mathbf{H}^H \mathbf{H} \right|$$

$$= \left| \mathbf{I} + \widehat{\mathbf{C}} \begin{bmatrix} \underline{0} & \underline{0} \\ 0 & (\mathbf{C}_2^H \mathbf{C}_2)^{-1} \end{bmatrix} \widehat{\mathbf{C}}^H \cdot \left[\mathbf{C}^{-H} \begin{bmatrix} \mathbf{\Lambda}_1 & 0 \\ 0 & \mathbf{\Lambda}_2 \end{bmatrix} \mathbf{C}^{-1} - \mathbf{I} \right] \right|$$

$$= \left| \mathbf{I} + \begin{bmatrix} \underline{0} & \underline{0} \\ \underline{0} & (\mathbf{C}_2^H \mathbf{C}_2)^{-1} \end{bmatrix} \cdot \left[\widehat{\mathbf{C}}^H \mathbf{C}^{-H} \begin{bmatrix} \mathbf{\Lambda}_1 & 0 \\ 0 & \mathbf{\Lambda}_2 \end{bmatrix} \mathbf{C}^{-1} \widehat{\mathbf{C}} - \widehat{\mathbf{C}}^H \widehat{\mathbf{C}} \right] \right|$$

$$= \left| \mathbf{I} + \begin{bmatrix} \underline{0} & \underline{0} \\ \underline{0} & (\mathbf{C}_2^H \mathbf{C}_2)^{-1} \end{bmatrix} \cdot \left[\begin{bmatrix} \mathbf{I} & \mathbf{N}^H \\ \underline{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{\Lambda}_1 & 0 \\ 0 & \mathbf{\Lambda}_2 \end{bmatrix} \begin{bmatrix} \mathbf{I} & \underline{0} \\ \mathbf{N} & \mathbf{I} \end{bmatrix} - \widehat{\mathbf{C}}^H \widehat{\mathbf{C}} \right] \right| \tag{71}$$

$$= \left| \mathbf{I} + \begin{bmatrix} \underline{0} & \underline{0} \\ \underline{0} & (\mathbf{C}_2^H \mathbf{C}_2)^{-1} \end{bmatrix} \cdot \left[\begin{bmatrix} \mathbf{\Lambda}_1 + \mathbf{N}^H \mathbf{\Lambda}_2 \mathbf{N} & \mathbf{N}^H \mathbf{\Lambda}_2 \\ \mathbf{\Lambda}_2 \mathbf{N} & \mathbf{\Lambda}_2 \end{bmatrix} - \widehat{\mathbf{C}}^H \widehat{\mathbf{C}} \right] \right|$$

$$= \left\| \begin{bmatrix} \mathbf{I} & \underline{\mathbf{0}} \\ (\mathbf{C}_2^H \mathbf{C}_2)^{-1} \mathbf{\Lambda}_2 \mathbf{N} & (\mathbf{C}_2^H \mathbf{C}_2)^{-1} \mathbf{\Lambda}_2 \end{bmatrix} \right\|$$

$$= \left| (\mathbf{C}_2^H \mathbf{C}_2)^{-1} \mathbf{\Lambda}_2 \right| = \left| (\mathbf{C}_2^H \mathbf{C}_2)^{-1} \right| \cdot |\mathbf{\Lambda}_2| \ , \tag{72}$$

where in (71), $\mathbf{N} = \left(\mathbf{C}_2^H \mathbf{P}_{\mathbf{C}_1}^\perp \mathbf{C}_2 \right)^{-1} \mathbf{C}_2^H \mathbf{P}_{\mathbf{C}_1}^\perp \mathbf{P}_{\mathbf{C}_2}^\perp \mathbf{C}_1$, and we used the fact that

$$\mathbf{C}^{-1} \widehat{\mathbf{C}} = \begin{bmatrix} (\mathbf{C}_1^H \mathbf{P}_{\mathbf{C}_2}^\perp \mathbf{C}_1)^{-1} \mathbf{C}_1^H \mathbf{P}_{\mathbf{C}_2}^\perp \\ (\mathbf{C}_2^H \mathbf{P}_{\mathbf{C}_1}^\perp \mathbf{C}_2)^{-1} \mathbf{C}_2^H \mathbf{P}_{\mathbf{C}_1}^\perp \end{bmatrix} \begin{bmatrix} \mathbf{P}_{\mathbf{C}_2}^\perp \mathbf{C}_1 & \mathbf{C}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \underline{\mathbf{0}} \\ \mathbf{N} & \mathbf{I} \end{bmatrix} .$$

Similarly, we have

$$\log \left| \mathbf{I} + \mathbf{S}^{\frac{1}{2}} \mathbf{P}_{\mathbf{C}_2}^\perp \mathbf{S}^{\frac{1}{2}} \mathbf{G}^H \mathbf{G} \right|$$

$$= \left| \mathbf{I} + \begin{bmatrix} (\mathbf{C}_1^H \mathbf{P}_{\mathbf{C}_2}^\perp \mathbf{C}_1)^{-1} & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \underline{\mathbf{0}} \end{bmatrix} \cdot \left[\begin{bmatrix} \mathbf{I} & \mathbf{N}^H \\ \underline{\mathbf{0}} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{I} & \underline{\mathbf{0}} \\ \mathbf{N} & \mathbf{I} \end{bmatrix} - \widehat{\mathbf{C}}^H \widehat{\mathbf{C}} \right] \right|$$

$$= \left| \mathbf{I} + \begin{bmatrix} (\mathbf{C}_1^H \mathbf{P}_{\mathbf{C}_2}^\perp \mathbf{C}_1)^{-1} & \underline{\mathbf{0}} \\ \underline{\mathbf{0}} & \underline{\mathbf{0}} \end{bmatrix} \cdot \left[\begin{bmatrix} \mathbf{I} + \mathbf{N}^H \mathbf{N} & \mathbf{N}^H \\ \mathbf{N} & \mathbf{I} \end{bmatrix} - \widehat{\mathbf{C}}^H \widehat{\mathbf{C}} \right] \right|$$

$$= \left\| \begin{bmatrix} (\mathbf{C}_1^H \mathbf{P}_{\mathbf{C}_2}^\perp \mathbf{C}_1)^{-1} (\mathbf{I} + \mathbf{N}^H \mathbf{N}) & (\mathbf{C}_1^H \mathbf{P}_{\mathbf{C}_2}^\perp \mathbf{C}_1)^{-1} \mathbf{N}^H \\ \underline{\mathbf{0}} & \mathbf{I} \end{bmatrix} \right\|$$

$$= \left| (\mathbf{C}_1^H \mathbf{P}_{\mathbf{C}_2}^\perp \mathbf{C}_1)^{-1} \right| \cdot \left| \mathbf{I} + \mathbf{N}^H \mathbf{N} \right| \ . \tag{73}$$

Subsituting (70), (72) and (73) in (69), we have $R_1 = \max(0, \log |\mathbf{\Lambda}_1| - \log \left| \mathbf{I} + \mathbf{N}^H \mathbf{N} \right|)$, which completes the proof.

APPENDIX E

PROOF OF LEMMA 2

We want to show that $\mathbf{W} \mathbf{H}^H \mathbf{H} \mathbf{W}$ and $\mathbf{W} \mathbf{G}^H \mathbf{G} \mathbf{W}$ commute, where $\mathbf{W} = (\mathbf{H}^H \mathbf{H} + \mathbf{G}^H \mathbf{G})^{-\frac{1}{2}}$ Let the invertible matrix $\widehat{\mathbf{C}}$ and diagonal matrix $\widehat{\mathbf{\Lambda}}$ respectively represent the generalized eigenvectors and eigenvalues of $\left(\mathbf{W} \mathbf{H}^H \mathbf{H} \mathbf{W} + \mathbf{I} \ ; \ \mathbf{W} \mathbf{G}^H \mathbf{G} \mathbf{W} + \mathbf{I} \right)$, so that

$$\widehat{\mathbf{C}}^H \left[\mathbf{W} \mathbf{H}^H \mathbf{H} \mathbf{W} + \mathbf{I} \right] \widehat{\mathbf{C}} = \widehat{\mathbf{\Lambda}} \tag{74}$$

$$\widehat{\mathbf{C}}^H \left[\mathbf{WG}^H \mathbf{GW} + \mathbf{I} \right] \widehat{\mathbf{C}} = \mathbf{I} \,. \tag{75}$$

Adding (74) and (75), we have

$$\widehat{\mathbf{C}}^H \left[\mathbf{W}(\mathbf{H}^H \mathbf{H} + \mathbf{G}^H \mathbf{G})\mathbf{W} + 2\,\mathbf{I} \right] \widehat{\mathbf{C}} = 3\,\widehat{\mathbf{C}}^H \widehat{\mathbf{C}} = (\widehat{\mathbf{\Lambda}} + \mathbf{I}) \,,$$

from which it results that $\widehat{\mathbf{C}}$ must be of the form [12]

$$\widehat{\mathbf{C}} = \frac{1}{\sqrt{3}} \, \mathbf{\Phi_w} (\widehat{\mathbf{\Lambda}} + \mathbf{I})^{\frac{1}{2}} \,, \tag{76}$$

where $\mathbf{\Phi_w}$ is an unknown unitary matrix. In the following, as we continue the proof, $\mathbf{\Phi_w}$ is characterized too.

Substituting (76) in (74) and (75), it is revealed that the unitary matrix $\mathbf{\Phi_w}$ represents the common set of eigenvectors for the matrices $\mathbf{WH}^H \mathbf{HW} + \mathbf{I}$ and $\mathbf{WG}^H \mathbf{GW} + \mathbf{I}$, and thus both matrices commute. In particular,

$$\mathbf{\Phi_w}^H \left[\mathbf{WH}^H \mathbf{HW} + \mathbf{I} \right] \mathbf{\Phi_w} = 3\,\widehat{\mathbf{\Lambda}}(\widehat{\mathbf{\Lambda}} + \mathbf{I})^{-1} = 3\,(\widehat{\mathbf{\Lambda}}^{-1} + \mathbf{I})^{-1}$$

$$\mathbf{\Phi_w}^H \left[\mathbf{WG}^H \mathbf{GW} + \mathbf{I} \right] \mathbf{\Phi_w} = 3\,(\widehat{\mathbf{\Lambda}} + \mathbf{I})^{-1} \,.$$

Consequently, $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are diagonal:

$$\begin{aligned}
\mathbf{\Phi_w}^H \, \mathbf{WH}^H \mathbf{HW} \, \mathbf{\Phi_w} &= 3\,(\widehat{\mathbf{\Lambda}}^{-1} + \mathbf{I})^{-1} - \mathbf{I} = \mathbf{\Sigma}_1 \\
\mathbf{\Phi_w}^H \, \mathbf{WG}^H \mathbf{GW} \, \mathbf{\Phi_w} &= 3\,(\widehat{\mathbf{\Lambda}} + \mathbf{I})^{-1} - \mathbf{I} = \mathbf{\Sigma}_2 \,.
\end{aligned} \tag{77}$$

It is interesting to note that, since $\mathbf{\Phi_w}^H \, \mathbf{WH}^H \mathbf{HW} \, \mathbf{\Phi_w} \succeq \underline{0}$ and $\mathbf{\Phi_w}^H \, \mathbf{WG}^H \mathbf{GW} \, \mathbf{\Phi_w} \succeq \underline{0}$, we have $\frac{1}{2}\mathbf{I} \preceq \widehat{\mathbf{\Lambda}} \preceq 2\,\mathbf{I}$.

APPENDIX F

PROOF OF LEMMA 3

We first consider the generalized eigenvalue decomposition for

$$\left(\mathbf{T_w}^H \mathbf{H}^H \mathbf{HT_w} + \mathbf{I} \,,\; \mathbf{T_w}^H \mathbf{G}^H \mathbf{GT_w} + \mathbf{I} \right) \,, \tag{78}$$

where $\mathbf{T_w}$ is given by (24) and

$$\overline{\mathbf{C}}_{\mathbf{w}}^{H} \left[\mathbf{T}_{\mathbf{w}}^{H} \mathbf{H}^{H} \mathbf{H} \mathbf{T_w} + \mathbf{I} \right] \overline{\mathbf{C}}_{\mathbf{w}} = \mathbf{\Lambda_w}$$

$$\overline{\mathbf{C}}_{\mathbf{w}}^{H} \left[\mathbf{T}_{\mathbf{w}}^{H} \mathbf{G}^{H} \mathbf{G} \mathbf{T_w} + \mathbf{I} \right] \overline{\mathbf{C}}_{\mathbf{w}} = \mathbf{I} .$$

Using (21), and noting that $\mathbf{\Phi_w}$ is unitary and $\mathbf{P}$ is diagonal, a straightforward calculation yields

$$\overline{\mathbf{C}}_{\mathbf{w}}^{H} \left[\mathbf{\Sigma}_1 \mathbf{P} + \mathbf{I} \right] \overline{\mathbf{C}}_{\mathbf{w}} = \mathbf{\Lambda_w}$$

$$\overline{\mathbf{C}}_{\mathbf{w}}^{H} \left[\mathbf{\Sigma}_2 \mathbf{P} + \mathbf{I} \right] \overline{\mathbf{C}}_{\mathbf{w}} = \mathbf{I} ,$$

where $\mathbf{\Sigma}_1$ and $\mathbf{\Sigma}_2$ are respectively (diagonal) eigenvalue matrices of $\mathbf{W}\mathbf{H}^{H}\mathbf{H}\mathbf{W}$ and $\mathbf{W}\mathbf{G}^{H}\mathbf{G}\mathbf{W}$, as given by (21). Thus, $\overline{\mathbf{C}}_{\mathbf{w}}$ is diagonal and is given by

$$\overline{\mathbf{C}}_{\mathbf{w}} = (\mathbf{\Sigma}_2 \mathbf{P} + \mathbf{I})^{-\frac{1}{2}} . \tag{79}$$

Consequently, we have $\mathbf{\Lambda_w} = (\mathbf{\Sigma}_2 \mathbf{P} + \mathbf{I})^{-1} (\mathbf{\Sigma}_1 \mathbf{P} + \mathbf{I})$.

Let σ_{1i}, σ_{2i} and p_i represent the ith diagonal elements of $\mathbf{\Sigma}_1$, $\mathbf{\Sigma}_2$ and $\mathbf{P}$, respectively. We note that for any p_i, $(1 + \sigma_{1i} p_i)/(1 + \sigma_{2i} p_i) > 1$ iff $\sigma_{1i} > \sigma_{2i}$. Thus, based on the argument that we made after Lemma 2, the first ρ diagonal elements of $\mathbf{\Lambda_w}$ represent generalized eigenvalues greater than 1. Letting

$$\mathbf{P} = \begin{bmatrix} \mathbf{P}_1 & \underline{0} \\ \underline{0} & \mathbf{P}_2 \end{bmatrix} \tag{80}$$

where $\mathbf{P}_1$ is $\rho \times \rho$ and $\mathbf{P}_2$ is $(n_t - \rho) \times (n_t - \rho)$, we have:

$$\mathbf{\Lambda_w} = \begin{bmatrix} (\mathbf{\Sigma}_{2\rho}\mathbf{P}_1 + \mathbf{I})^{-1}(\mathbf{\Sigma}_{1\rho}\mathbf{P}_1 + \mathbf{I}) & \underline{0} \\ \underline{0} & (\mathbf{\Sigma}_{2\bar{\rho}}\mathbf{P}_2 + \mathbf{I})^{-1}(\mathbf{\Sigma}_{1\bar{\rho}}\mathbf{P}_2 + \mathbf{I}) \end{bmatrix} = \begin{bmatrix} \mathbf{\Lambda}_{1\mathbf{w}} & \underline{0} \\ \underline{0} & \mathbf{\Lambda}_{2\mathbf{w}} \end{bmatrix} , \tag{81}$$

where $\mathbf{\Sigma}_{i\rho}$ and $\mathbf{\Sigma}_{i\bar{\rho}}$ ($i = 1, 2$) are given by (22). Consequently, (79) can be rewritten as

$$\overline{\mathbf{C}}_{\mathbf{w}} = [\overline{\mathbf{C}}_{1\mathbf{w}} \quad \overline{\mathbf{C}}_{2\mathbf{w}}] = \begin{bmatrix} (\mathbf{\Sigma}_{2\rho}\mathbf{P}_1 + \mathbf{I})^{-\frac{1}{2}} & \underline{0} \\ \underline{0} & (\mathbf{\Sigma}_{2\bar{\rho}}\mathbf{P}_2 + \mathbf{I})^{-\frac{1}{2}} \end{bmatrix} . \tag{82}$$

From the argument before Lemma 3, for any diagonal

$\mathbf{P} \succeq \underline{0}$, linear precoding is an optimal solution for the BC under the matrix power constraint $\mathbf{S_w} = \mathbf{T_w} \mathbf{T_w^H}$, where $\mathbf{T_w}$ is given by (24). More precisely, from Theorem 1, $X = V_1 + V_2$ is optimal, where V_1 and V_2 are independent Gaussian precoders, respectively corresponding to W_1 and W_2 with zero means and covariance matrices $\mathbf{K_{tw}^*}$ and $\mathbf{S_w} - \mathbf{K_{tw}^*}$, where $\mathbf{K_{tw}^*}$ is given by

$$\mathbf{K_{tw}^*} = \mathbf{S_w^{\frac{1}{2}}} \mathbf{C_w} \begin{bmatrix} (\mathbf{C_{1w}^H} \mathbf{C_{1w}})^{-1} & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \mathbf{C_w^H} \mathbf{S_w^{\frac{1}{2}}} \tag{83}$$

and $\mathbf{C_w}$ is the generalized eigenvector matrix for

$$\left(\mathbf{S_w^{\frac{1}{2}}} \mathbf{H^H} \mathbf{H} \mathbf{S_w^{\frac{1}{2}}} + \mathbf{I} \, , \, \mathbf{S_w^{\frac{1}{2}}} \mathbf{G^H} \mathbf{G} \mathbf{S_w^{\frac{1}{2}}} + \mathbf{I} \right) . \tag{84}$$

We note that there exists a unitary matrix $\boldsymbol{\Psi}$ for which $\mathbf{S_w^{\frac{1}{2}}} = \mathbf{T_w} \boldsymbol{\Psi}^H$ [12], where $\mathbf{S_w} = \mathbf{T_w} \mathbf{T_w^H}$. We also note that, from Remark 4, $\mathbf{C_w} = \boldsymbol{\Psi} \overline{\mathbf{C}}_\mathbf{w}$ and $\mathbf{C_w^H} \mathbf{C_w} = \overline{\mathbf{C}}_\mathbf{w}^H \overline{\mathbf{C}}_\mathbf{w}$. Thus, $\mathbf{K_{tw}^*}$ can be rewritten as

$$\mathbf{K_{tw}^*} = \mathbf{T_w} \overline{\mathbf{C}}_\mathbf{w} \begin{bmatrix} (\overline{\mathbf{C}}_{1\mathbf{w}}^H \overline{\mathbf{C}}_{1\mathbf{w}})^{-1} & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \overline{\mathbf{C}}_\mathbf{w}^H \mathbf{T_w^H} = \mathbf{T_w} \begin{bmatrix} \mathbf{I} & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \mathbf{T_w^H} \tag{85}$$

$$= \mathbf{W} \boldsymbol{\Phi}_\mathbf{w} \mathbf{P}^{\frac{1}{2}} \begin{bmatrix} \mathbf{I} & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \mathbf{P}^{\frac{1}{2}} \boldsymbol{\Phi}_\mathbf{w}^H \mathbf{W} = \mathbf{W} \boldsymbol{\Phi}_\mathbf{w} \begin{bmatrix} \mathbf{P_1} & \underline{0} \\ \underline{0} & \underline{0} \end{bmatrix} \boldsymbol{\Phi}_\mathbf{w}^H \mathbf{W} \, , \tag{86}$$

where (85) comes from (82), and (86) comes from (80). Consequently, $\mathbf{S_w} - \mathbf{K_{tw}^*}$ can be written as

$$\mathbf{S_w} - \mathbf{K_{tw}^*} = \mathbf{T_w} \mathbf{T_w^H} - \mathbf{K_{tw}^*} = \mathbf{W} \boldsymbol{\Phi}_\mathbf{w} \mathbf{P} \boldsymbol{\Phi}_\mathbf{w}^H \mathbf{W} - \mathbf{K_{tw}^*}$$

$$= \mathbf{W} \boldsymbol{\Phi}_\mathbf{w} \begin{bmatrix} \underline{0} & \underline{0} \\ \underline{0} & \mathbf{P_2} \end{bmatrix} \boldsymbol{\Phi}_\mathbf{w}^H \mathbf{W} \, . \tag{87}$$

From (86) and (87), under the matrix power constraint $\mathbf{S_w}$ given by (23), the optimal linear precoding is $X = V_1 + V_2$, where precoding signals V_1 and V_2 are independent Gaussian vectors with zero means and covariance matrices given by (86) and (87), respectively. Alternatively, the optimal precoder can be represented as $X = \mathbf{W} \boldsymbol{\Phi}_\mathbf{w} \begin{bmatrix} V_1' \\ V_2' \end{bmatrix}$, where precoding signals V_1'

and V_2' are independent Gaussian vectors with zero means and diagonal covariance matrices respectively given by $\mathbf{P}_1$ and $\mathbf{P}_2$. In both cases $\mathbb{E}\{XX^H\} = \mathbf{S_w}$, and the same secrecy rate region is achieved.

REFERENCES

[1] A. Wyner, "The wire-tap channel," *Bell. Syst. Tech. J.*, vol. 54, no. 8, pp. 1355-1387, Jan. 1975.

[2] R. Liu, I. Maric, P. Spasojevic, and R. D. Yates, "Discrete memoryless interference and broadcast channels with confidential messages: Secrecy rate regions," *IEEE Trans. Inf. Theory*, vol. 54, no. 6, pp. 2493-2512, June 2008.

[3] R. Liu and H. V. Poor, "Secrecy capacity region of a multiple-antenna Gaussian broadcast channel with confidential messages," *IEEE Trans. Inf. Theory*, vol. 55, no. 3, pp. 1235-1249, Mar. 2009.

[4] R. Liu, T. Liu, H. V. Poor, and S. Shamai, "Multiple-input multiple-output Gaussian broadcast channels with confidential messages," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4215-4227, 2010.

[5] Q. H. Spencer, A. L. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Processing*, vol. 52, no. 2, pp. 461-471, Feb. 2004.

[6] T. Yoo and A. Goldsmith, "On the optimality of multi-antenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Select. Areas Commun.*, Special issue on 4G Wireless Systems, vol. 24, no. 3, pp.528-541, Mar. 2006.

[7] A. Wiesel, Y. Eldar, and S. Shamai, "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Trans. Signal Processing*, vol. 54, no. 1, pp. 161-176, Jan. 2006.

[8] S. Ali. A. Fakoorian and A. L. Swindlehurst, "Dirty Paper Coding Versus Linear GSVD-Based Precoding in MIMO Broadcast Channel with Confidential Messages," in *Proc. IEEE Globecom*, Dec. 2011.

[9] A. Khisti and G. Wornell, "Secure transmission with multiple antennas II: The MIMOME wiretap channel," *IEEE Trans. Inf. Theory*, vol. 56, no. 11, pp. 5515-5532, 2010.

[10] S. Ali. A. Fakoorian and A. L. Swindlehurst, "Optimal Power Allocation for the GSVD based MIMO Gaussian Wiretap Channel," in *ISIT*, July 2012.

[11] R. Bustin, R. Liu, H. V. Poor, and S. Shamai (Shitz), "A MMSE approach to the secrecy capacity of the MIMO Gaussian wiretap channel," *EURASIP Journal on Wireless Comm. and Net.*, vol. 2009, Article ID 370970, 8 pages, 2009.

[12] R. A. Horn and C. R. Johnson, *Matrix Analysis*, University Press, Cambridge, UK, 1985.

[13] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936-3964, 2006.

[14] D. Nion, "A Tensor Framework for Nonunitary Joint Block Diagonalization," *IEEE Trans. Signal Processing*, vol. 59, no. 10, pp. 4585-4594, Oct. 2011.

[15] S. A. A. Fakoorian and A. L. Swindlehurst, "MIMO interference channel with confidential messages: achievable secrecy rates and beamforming design," *IEEE Trans. on Inf. Forensics and Security*, vol. 6, no. 3, Sep. 2011.

[16] J. Lee, and N. Jindal, "High SNR Analysis for MIMO Broadcast Channels: Dirty Paper Coding Versus Linear Precoding," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4787-4792, Dec. 2007.